

Accession # 100453

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Name Moshir Accession # 67746 Date 8/5
 Accession # 1648 Phone Number 308-2926 Serial Number 091851,410
 Mail Box and Bag Room Location _____ Results Wanted Preferred Type PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need

Provide a detailed description of the search topic and describe in detail the information and subject matter to be searched. Use the following guidelines: keywords, synonyms, acronyms, and nomenclature numbers, and combine with the concept or subject of the invention. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. Please underline any key phrases, brief pertinent claims, and abstract.

Title of Invention _____

Inventors, producer, or other names _____

Earliest Priority Filing Date _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Search seq 8 str & oligo

STAFF USE ONLY

Searcher	Type of Search	Vendors and cost where applicable
<u>D. Schweitzer</u>	NA Sequence # _____	Cost _____
Searcher Phone # <u>308-4292</u>	AA Sequence # <u>2</u>	Dialog _____
Searcher Accession # <u>CM1 6A03</u>	Structure # _____	Disc # _____
Searcher Initials <u>DS</u>	Biography # _____	_____
Searcher Date <u>8/5</u>	Citation _____	_____
Searcher Initials & Phone # <u>13</u>	Full text _____	Sequence Systems <u>Com pay</u>
Searcher Step # _____	Parent Name _____	_____
Searcher Initials <u>6</u>	_____	_____





STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number

TO: Mary Mosher
Location: cm-1/9a17/8e12
Art Unit: 1648
Tuesday, August 05, 2003

Case Serial Number: 09/851410

From: David Schreiber
Location: Biotech-Chem Library
CM1-6A03
Phone: 308-4292

david.schreiber@uspto.gov

Search Notes

Search seq 8 std and oligo



reference version 5.1.6

1977 *Corvidae*: 8007 - 8561 (5) *Phylloscopus*

M protein protein search, using sw model

born on: August 5, 2003, 09:29:48 ; Search time 47 seconds

(without alignments)

2228.926 Million cell updates/sec

US 8,141,410A 8

1. The first step is to identify the key components of the system. This involves understanding the hardware and software involved, as well as the data flow and the roles of the various components.

1 MRRRRLILLIIIMFIIIMFIIIPAP
.....GCTVATTCPLKMFVAVKTEEL 660

1. Introduction

$$\hat{\theta}^{(1)}(y) = \hat{\theta}^{(0)}(y) + \hat{\theta}^{(1)}(x)$$

Source: 197864, 158726573, residues

1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 2680, 26

1. *Office Memorandum*

[illegible]

Maximum length: 2900000000

[illegible]

```

A_Geneset_19.Jun03.*
1: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1980.DAT.*
2: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1981.DAT.*
3: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1982.DAT.*
4: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1983.DAT.*
5: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1984.DAT.*
6: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1985.DAT.*
7: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1986.DAT.*
8: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1987.DAT.*
9: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1988.DAT.*
10: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1989.DAT.*
11: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1990.DAT.*
12: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1991.DAT.*
13: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1992.DAT.*
14: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1993.DAT.*
15: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1994.DAT.*
16: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1995.DAT.*
17: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1996.DAT.*
18: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1997.DAT.*
19: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1998.DAT.*
20: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA1999.DAT.*
21: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA2000.DAT.*
22: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA2001.DAT.*
23: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA2002.DAT.*
24: /S/HSL/qqdata/a/qgeneset/qgenesetp-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Isolate No.	*			ID	Description
	Score	Length	Match		
1	66.0	100.0	66.0	AA848786	HV GP2 protein.
2	66.0	100.0	66.0	AA849406	Burma strain HV 0
3	66.0	100.0	66.0	AA866089	Hepatitis E virus
4	66.0	100.0	66.0	AA835426	Hepatitis E virus
5	66.0	100.0	66.0	AA860197	Protein encoded by
6	66.0	100.0	66.0	AA893489	Human HV ORF 2 pr
7	66.0	100.0	66.0	AA824120	Hepatitis E virus
8	66.0	100.0	66.0	AA862524	HV Burma strain v
9	66.0	100.0	66.0	AA815699	Hepatitis E virus


```

A* AAB41120;
XX 29-JAN-2001 (first entry)
DE Hepatitis E virus Burma strain HEF2 protein sequence REF ID: NEB-R
KW Hepatitis E virus; HEV; non-A/non-B viral hepatitis; viral protein;
KW immunological; diagnosis; hepatitis; infection; identification;
KW diagnostic; immunoreactive; hepatitis; HEV; identification; antigen;
KW vaccine; antiviral; antigenic; antibody; antigen.
XX
OS Hepatitis E virus.
XX
PN US6120988-A.
XX
PD 19-SEP-2000.
XX
PF 07-JUN-1995; 95US-0478507.
XX
PR 25-JUL-1994; 94US-0279823.
PR 05-APR-1991; 91US 0681078.
PR 17-JUN-1988; 88US-0208997.
PR 11-APR-1989; 88US-0336672.
PR 16-JUN-1989; 89US-0367486.
PR 13-OCT-1989; 89US 0429921.
PR 05-APR-1990; 90US-0505888.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
PA (USG-) US GOVERNMENT.
XX
PI Yartsevich DN, Krawczynski K; Fry ZF, Bradley DW, Tarr A, Reyes DF;
XX WPL; 1990; 94712-66.
DR N-PSDB: AAA99259.
XX
PI Identifying recombinant antigen immunoreactive with antibody induced by
PI hepatitis E virus (HEV) for detecting HEV infection; epitopes
PI immunoreacting a polypeptide from an HEV genome with an HEV-positive
PI antiserum.
XX
PS Claim 6; Column 63-68; 46pp; English.
XX
CC The present invention describes a method for identifying a recombinant
CC antigen immunoreactive with a hepatitis E virus (HEV)-induced antibody.
CC The method comprises producing a polypeptide derived from an HEV genome,
CC immunoreacting the polypeptide with an HEV-positive antiserum and
CC selecting the polypeptide as a recombinant antigen if the polypeptide
CC reacts with the HEV-positive antiserum. The method is useful for
CC identifying recombinant antigen immunoreactive with antibody induced by
CC HEV. The enterically transmitted non-A/non-B hepatitis (ET-NANB (also
CC known as HEV))-specific fragments are useful for identifying ET-NANB-
CC derived cDNAs, which contain additional sequence information, as probes
CC for detecting ET-NANB viral genomic material in a patient sample, for
CC the synthesis of polypeptides for use in immunoassays, and for
CC identifying similar antigenic regions encoded by related viral strains,
CC e.g. Burmese strain. The antigens are especially useful in the
CC preparation of vaccine against ET-NANB infection. These antigens may
CC further be used to prepare antibodies to ET-NANB virus particles for
CC use directly as antiviral agents, and to produce antiserum designed for
CC pre or post exposure prophylaxis. The present sequence represents a
CC specifically claimed HEV Burma strain protein sequence for use in the
CC present invention.
XX
SQ Sequence 660 AA:
Query Match 100.0%; Score 660; DB 21; Length 660;
Best local Similarity 100.0%; Pred No. 0;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Caps 0;
XX 1 MFPPPTTLLTMTPLMIPADPPQSPSPPPSPSPSSGSGGCPWTDVWVSQPPAIPYIPTN 60
DB 1 MFPPPTTLLTMTPLMIPADPPQSPSPPPSPSPSSGSGGCPWTDVWVSQPPAIPYIPTN 60

```

P1 Reyes JR, Yarbough JW, Krawczynski KZ, Lam AW, Fry KL.
 XX WFL: 2001-04-23/05-26.
 DB N-15DB: AAL50486.
 XX New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)
 P1 hepatitis viral agent, useful in diagnosing infection by an enterically
 P1 transmitted agent (e.g. ET-NANB virus), as well as in vaccine
 P1 production.
 XX Disclosure: Columns 64-68: 45pp; English.
 PS the invention relates to an isolated DNA comprising the genome of an
 XX enterically transmitted non-A/non-B (ET-NANB) viral hepatitis agent (also
 CC referred as HEV). The DNA sequences or their fragments are useful in
 CC preparing ET-NANB viral proteins and as probes for virus detection.
 CC these are particularly useful in diagnosing infection by an enterically
 CC transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
 CC the present sequence represents a open? protein encoded by a ET-NANB viral
 CC DNA sequence from HEV-Burma strain
 XX
 SS Sequence: 650 AA:

Query Match 100.0%; Score 660; Ig 22; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPRPIELLLMLLPIAPPGQSGRRGSGCGFWGDRVUSQPAIPYIHTN 60
 DB 1 MPPRPIELLLMLLPIAPPGQSGRRGSGCGFWGDRVUSQPAIPYIHTN 60
 QY 61 PRAPVTAAGAGPVPGPAPPLGSAWPDQAGPAAVASRRPPTTAGAAPTAVAPADTP 120
 DB 61 PRAPVTAAGAGPVPGPAPPLGSAWPDQAGPAAVASRRPPTTAGAAPTAVAPADTP 120
 QY 121 PPDVLSKAILRQYNIISLTSSVATGTNI VLYAAPLSPLLPLQDGTNTHIMAPEAS 180
 DB 121 PPDVLSKAILRQYNIISLTSSVATGTNI VLYAAPLSPLLPLQDGTNTHIMAPEAS 180
 QY 181 NYAQYKVAHATIKYRPLVNAVGGYALISFWPQTITPTTSVDMNSITSDVRLVQPGI 240
 DB 181 NYAQYKVAHATIKYRPLVNAVGGYALISFWPQTITPTTSVDMNSITSDVRLVQPGI 240
 QY 241 ASSELVPSPPHYRQGWPSVFTSGVAFFEATSGVLMLCHGSLVNSYTNTPYTGALCL 300
 DB 241 ASSELVPSPPHYRQGWPSVFTSGVAFFEATSGVLMLCHGSLVNSYTNTPYTGALCL 300
 QY 301 DALALEPRLNLTPTNTVSPYSSSTAPHPPTPPADPTAFITTTAATFPKPDLYFTSTNG 360
 DB 301 DALALEPRLNLTPTNTVSPYSSSTAPHPPTPPADPTAFITTTAATFPKPDLYFTSTNG 360
 QY 361 VKEIGGIALTLFNLADILGGLPELTISSAGGGLPYSPVVSANGEPYVLYTSVHNAQ 420
 DB 361 VKEIGGIALTLFNLADILGGLPELTISSAGGGLPYSPVVSANGEPYVLYTSVHNAQ 420
 QY 421 QERGLAETRLVLEDSERVVLCQVQKHEQDPTTSPAFRRFPFVSLEANTWMLTSLTAIFY 480
 DB 421 QERGLAETRLVLEDSERVVLCQVQKHEQDPTTSPAFRRFPFVSLEANTWMLTSLTAIFY 480
 QY 481 DQSTYSSSTGPVVSIVSLVNVAPGAAVAPSLDWIKVLLPDPPLSTLQYSKTFEVL 540
 DB 481 DQSTYSSSTGPVVSIVSLVNVAPGAAVAPSLDWIKVLLPDPPLSTLQYSKTFEVL 540
 QY 541 LPRGLSEWAGTFAWATVRYSTAAAGLVNNAQHEVALVLTSLAGAGVLSAVAV 600
 DB 541 LPRGLSEWAGTFAWATVRYSTAAAGLVNNAQHEVALVLTSLAGAGVLSAVAV 600
 QY 601 LAHPSALALFDTLLYPAARHTFDYPTFPAGLQYQAFQSTVAELQRLMKVGTTEE 660
 DB 601 LAHPSALALFDTLLYPAARHTFDYPTFPAGLQYQAFQSTVAELQRLMKVGTTEE 660

RESULT 9

AAL5699
 ID AAL5699 standard: Protein: 650 AA
 XX
 AC AAL5699;
 XX
 XX 08-NOV-2002 (first entry)
 XX Hepatitis E virus (Burma strain) REF2 protein
 DE
 XX HEV, enterically transmitted nonA/nonB hepatitis viral agent;
 KW Burma strain; bile; REF2;
 XX
 OS Hepatitis E virus;
 XX
 XX US6379891 BL;
 XX 30 APR 2002;
 XX 19 APR 2002; 200303 053427;
 XX
 PR 25-JUL-1994; 9405-0279827;
 PR 07-JUN-1995; 9405-0478502;
 PR 15-APR-1991; 9405-0681078;
 PR 17-JUN-1989; 9405-0208997;
 PR 11-APR-1989; 8405-0346672;
 PR 16-JUN-1989; 8405-0367486;
 PR 13-OCT-1989; 8405-0420925;
 PR 05-APR-1990; 9405-0505886;
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES;
 PA (GENE-) GENELABS TECHNOLIGIES INC;
 PA
 XX Reyes JR, Yarbough JW, Krawczynski KZ, Lam AW, Fry KL;
 XX
 DB WFL: 2002-51727755;
 DB N-15DB: AAL50486;
 XX
 PT Detecting the presence of enterically transmitted nonA/nonB hepatitis
 PT viral (HEV) agents in bile samples from infected humans and monkeys
 XX using polymerase chain reaction;
 PS Disclosure: Column 19-24: 6pp; English.
 XX The invention comprises a method for detecting the presence of
 CC enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a
 CC sample and isolating HEV agents or nucleic acid fragments produced by the
 CC agent. The method utilizes PCR using bile from a human or cynomolgus
 CC monkey actively infected with HEV as a source of the agent. The method of
 CC the invention is used for detecting the presence of a viral agent in a
 CC sample of cultured cells infected with the agent and isolating
 CC enterically transmitted nonA/nonB HEV agents or nucleic acid fragments
 CC produced by the agent. The present amino acid sequence represents the
 CC protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated
 CC from a Burmese strain of the Hepatitis E virus

XX Sequence: 650 AA:

Query Match 100.0%; Score 660; Ig 22; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPRPIELLLMLLPIAPPGQSGRRGSGCGFWGDRVUSQPAIPYIHTN 60
 DB 1 MPPRPIELLLMLLPIAPPGQSGRRGSGCGFWGDRVUSQPAIPYIHTN 60
 QY 61 PRAPVTAAGAGPVPGPAPPLGSAWPDQAGPAAVASRRPPTTAGAAPTAVAPADTP 120
 DB 61 PRAPVTAAGAGPVPGPAPPLGSAWPDQAGPAAVASRRPPTTAGAAPTAVAPADTP 120
 QY 121 PPDVLSKAILRQYNIISLTSSVATGTNI VLYAAPLSPLLPLQDGTNTHIMAPEAS 180
 DB 121 PPDVLSKAILRQYNIISLTSSVATGTNI VLYAAPLSPLLPLQDGTNTHIMAPEAS 180

QY 181 NYAGYVPAPATIPYPIPVNAVAGYATISISFWGQTITITISVMNSITSHWRIIVQPTI 240
 DB 181 NYAGYVPAPATIPYPIPVNAVAGYATISISFWGQTITITISVMNSITSHWRIIVQPTI 240
 QY 241 ASFLVTPSERIHYPNQGWPSVETSCVAEEFATSGIVMLCTHGSLSVNSYNTIYTSALGLL 300
 DB 241 ASFLVTPSERIHYPNQGWPSVETSCVAEEFATSGIVMLCTHGSLSVNSYNTIYTSALGLL 300
 QY 301 DPALELEPNNITPGNTNTRVSRYSSTAFHLPKCAQTAFITTAATPEMKRIYTSYNSG 360
 DB 301 DPALELEPNNITPGNTNTRVSRYSSTAFHLPKCAQTAFITTAATPEMKRIYTSYNSG 360
 QY 361 VGEIGRGIALTEPNIAPTEIIGRLPTELISAGGQLEYSREVSANGPEPTVKLYTSVENAQ 420
 DB 361 VGEIGRGIALTEPNIAPTEIIGRLPTELISAGGQLEYSREVSANGPEPTVKLYTSVENAQ 420
 QY 421 QERGIATPHRIIGGESHWVLDYUNQHEGRTFTSHVATSPPSVULKANIWIWLSIAAAY 480
 DB 421 QERGIATPHRIIGGESHWVLDYUNQHEGRTFTSHVATSPPSVULKANIWIWLSIAAAY 480
 QY 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWIEVLDEKPLSTIIGSKTPEVLP 540
 DB 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWIEVLDEKPLSTIIGSKTPEVLP 540
 QY 541 LPKLSLSEWAGTTKAGYPYNYNTTASQQLLVENAAHHPVAISTYTSIGAGDVSISAVAV 600
 DB 541 LPKLSLSEWAGTTKAGYPYNYNTTASQQLLVENAAHHPVAISTYTSIGAGDVSISAVAV 600
 QY 601 LAPHSALELELLETPAPAHTEPTEPPEPPIGIGQSTVAEIQPKMKVCKTPEI 660
 DB 601 LAPHSALELELLETPAPAHTEPTEPPEPPIGIGQSTVAEIQPKMKVCKTPEI 660

RESULT 10

AAW76369

ID AAW76369 standard; Protein: 660 AA.

XX AAW76369;

DI 03-DEC-1998 (first entry)

DE Hepatitis E virus hollow particle protein #2.

XX Hepatitis E virus hollow particle protein, virus, antibody, detection, immunodassay.

PW infection

OS Hepatitis virus.

XX JF1024383-A

XX 08-SEP 1998

PF 28-FEB-1997; 97JP-0062445.

XX 28-FEB-1997; 97JP-0062445

XX (DENK-) DENKA SEIKEN KK

PA (KORU-) KOKUTSU YAKU EISEI KENKYUSHO.

XX WPI: 1998-535037/46.

DR N-PSDR; AAW61688.

XX Hepatitis E virus hollow particle polypeptide(s) and nucleic acids.
 PT encoding it useful for more accurate detection of HEV in samples,
 PT using immunoassays and nucleic acid hybridisation.

PS Claim 13; Page 24-26; 29pp; Japanese.

XX This sequence represents a Hepatitis E viral hollow particle protein.
 CC This polypeptide can be used to raise antibodies to detect HEV infection
 CC in samples, e.g. by immuno assay based techniques, and the nucleic acid
 CC can be used for the same in nucleic acid hybridisation assays. The
 CC polypeptides and nucleic acids allow more accurate detection of HEV than

CC previously possible.

XX Sequence 660 AA;

Query Match 84.7%; Score 559; DB 19; Length 660;
 Best local Similarity 99.8%; Pred. No. 0;
 Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPPEFTILILMLPLMIPATPTTQFSTPEFPSSSSGGGCTWGTAVAGTFLITVIMV 60
 DB 1 MPPEFTILILMLPLMIPATPTTQFSTPEFPSSSSGGGCTWGTAVAGTFLITVIMV 60
 QY 61 PFAPFTVAAATATPPVSGQAPMIGSAMPATGAPPAVACPPRPDPAFAHFAVAHAPDP 120
 DB 61 PFAPFTVAAATATPPVSGQAPMIGSAMPATGAPPAVACPPRPDPAFAHFAVAHAPDP 120
 QY 121 FVLVVSERGALLEELVYMLSTLTSVATGTNIVLYAAPLTELLELQTTTHIMAEAS 180
 DB 121 FVLVVSERGALLEELVYMLSTLTSVATGTNIVLYAAPLTELLELQTTTHIMAEAS 180
 QY 181 NYAGYVPAPATIPYPIPVNAVAGYATISISFWGQTITITISVMNSITSHWRIIVQPTI 240
 DB 181 NYAGYVPAPATIPYPIPVNAVAGYATISISFWGQTITITISVMNSITSHWRIIVQPTI 240
 QY 241 ASFLVTPSERIHYPNQGWPSVETSCVAEEFATSGIVMLCTHGSLSVNSYNTIYTSALGLL 300
 DB 241 ASFLVTPSERIHYPNQGWPSVETSCVAEEFATSGIVMLCTHGSLSVNSYNTIYTSALGLL 300
 QY 301 DPALELEPNNITPGNTNTRVSRYSSTAFHLPKCAQTAFITTAATPEMKRIYTSYNSG 360
 DB 301 DPALELEPNNITPGNTNTRVSRYSSTAFHLPKCAQTAFITTAATPEMKRIYTSYNSG 360
 QY 361 VGEIGRGIALTEPNIAPTEIIGRLPTELISAGGQLEYSREVSANGPEPTVKLYTSVENAQ 420
 DB 361 VGEIGRGIALTEPNIAPTEIIGRLPTELISAGGQLEYSREVSANGPEPTVKLYTSVENAQ 420
 QY 421 QERGIATPHRIIGGESHWVLDYUNQHEGRTFTSHVATSPPSVULKANIWIWLSIAAAY 480
 DB 421 QERGIATPHRIIGGESHWVLDYUNQHEGRTFTSHVATSPPSVULKANIWIWLSIAAAY 480
 QY 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWIEVLDEKPLSTIIGSKTPEVLP 540
 DB 481 DQSTYGSSTGPPVVSSTLVNATAGAVAPSLIMWIEVLDEKPLSTIIGSKTPEVLP 540
 QY 541 LPKLSLSEWAGTTKAGYPYNYNTTASQQLLVENAAHHPVAISTYTSIGAGDVSISAVAV 600
 DB 541 LPKLSLSEWAGTTKAGYPYNYNTTASQQLLVENAAHHPVAISTYTSIGAGDVSISAVAV 600
 QY 601 LAPHSALELELLETPAPAHTEPTEPPEPPIGIGQSTVAEIQPKMKVCKTPEI 660
 DB 601 LAPHSALELELLETPAPAHTEPTEPPEPPIGIGQSTVAEIQPKMKVCKTPEI 660

RESULT 11

AAW71210

ID AAW71210 standard; Protein: 660 AA.

XX AAW71210;

XX 25-MAR-2003 (updated)

DT 30-OCT-1998 (first entry)

XX Protein encoded by ORF 3 of the human isolate of HT-NANP.
 XX Encoded, transmitted, and sequenced by the Japanese
 KW HEV; HT-NANP; detection; vaccine.
 XX Hepatitis virus.

XX Key location/qualifiers

FT Misc-difference 496

PT /note- "not specified"

XX

KW HEV: externally transmitted non-African hepatitis virus; vaccine;
 KW diagnostic antigen; Spodoptera frugiperda; 51% insect;
 KW baculovirus; capsid.
 XX
 XX Hepatitis E virus Burma strain.
 XX W09612697 A2
 XX
 XX 02 MAY 1996.
 XX 23 OCT 1995; 95W0514703.
 XX 14 OCT 1995; 950S 0542644.
 XX 24 OCT 1994; 940S-0427952.
 XX (GENE) GENEJARS TECHNOLIGIES INC.
 XX
 XX Fucst TR, McAtee CP, Yarbough PO, Zhang Y.
 XX W01; 1996-230608/23.
 XX
 XX Hepatitis E virus (HEV) antigens derived from opp 2 - useful as
 XX diagnostic reagents for determining HEV infection and in vaccines
 XX
 XX Claim 5; Page 92-93; 125pp; English.
 XX
 XX DNA (AA127109) coding for the 62K antigen (AA96091) of hepatitis E
 XX virus (HEV) Burma strain capsid protein was cloned into baculovirus
 XX expression vector pBaculatt and recombinant 62K was expressed
 XX in Sf9 insect cells. High levels of expression were obtained and
 XX the recombinant 62K was obtained in over 95% purity. However, C-terminal
 XX processing resulted in the deletion of 9 or 23 amino acids from 62K,
 XX giving 2 related polypeptide species (AAR96101 and AAR96103). Similar
 XX results were obtained with HEV Mexico 62K antigen (see also AAR96102 and
 XX AAR96104). Recombinant 62K represents an improved antigen, in
 XX comparison to bacterial expressed proteins, for use in HEV
 XX diagnostic assays, and also has excellent immunogenic properties.
 XX Sequence 540 AA:

Query Match 81.8%; Score 540; DB 17; Length 540;
 Best Local Similarity 100.0%; Prod. No. 0;
 Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 0Y 112 AVAFAHPPPPVGVLSGCAILKRYNISTSPSSVAFGINVLYAARLSPFLPLGATN 171
 0b 1 AVAFAHPPPPVGVLSGCAILKRYNISTSPSSVAFGINVLYAARLSPFLPLGATN 69
 0Y 172 THIMATGAVYACVAVAPATRYEPLVHAWGVNA:STGWP:TTTTTAVMREITST 271
 0b 61 THIMATGAVYACVAVAPATRYEPLVHAWGVNA:STGWP:TTTTTAVMREITST 129
 0Y 242 99H1W2PCLASGLVLPSEKLYHPCWPSVPLEVAPEFATSLVMLCIGESLVNSYNT 291
 0b 121 VHLWPGHAGSLVLPSEKLYHPCWPSVPLEVAPEFATSLVMLCIGESLVNSYNT 189
 0Y 272 PVTALGLLEDALEEFNTEFNINRVSSYSTASHRLRFGAGTIAELTTTAATPFMR 351
 0b 101 PVTALGLLEDALEEFNTEFNINRVSSYSTASHRLRFGAGTIAELTTTAATPFMR 249
 0Y 312 PLYPSTSPWVSTPRGTAALFNINATILGCLPTELLSSAGQPLSPRVVSSANGPTV 411
 0b 241 PLYPSTSPWVSTPRGTAALFNINATILGCLPTELLSSAGQPLSPRVVSSANGPTV 309
 0Y 474 LITAVERA:QWKAATHR:ELAG:SAVVLELQ:QH:ELKPTT:SG:AT:SP:SG:KARI:V: 471
 0b 601 LITAVERA:QWKAATHR:ELAG:SAVVLELQ:QH:ELKPTT:SG:AT:SP:SG:KARI:V: 369
 0Y 474 WLSH:TAAPV:AG:SYGSS:GPVYV:DSVTLVNVAG:AG:AVAS:KMT:KVI:LP:K:PLST:QQ 541
 0b 361 WLSH:TAAPV:AG:SYGSS:GPVYV:DSVTLVNVAG:AG:AVAS:KMT:KVI:LP:K:PLST:QQ 429
 0Y 542 TSE:TTV:V:LL:DEP:STW:AT:TEAT:TE:TH:HTA:PL:GL:V:V:NA:AG:HV:V:STTT:TC:AG 591

0A 112 AVAFAHPPPPVGVLSGCAILKRYNISTSPSSVAFGINVLYAARLSPFLPLGATN 480
 0b 1 AVAFAHPPPPVGVLSGCAILKRYNISTSPSSVAFGINVLYAARLSPFLPLGATN 69
 0Y 172 THIMATGAVYACVAVAPATRYEPLVHAWGVNA:STGWP:TTTTTAVMREITST 271
 0b 61 THIMATGAVYACVAVAPATRYEPLVHAWGVNA:STGWP:TTTTTAVMREITST 129
 0Y 242 99H1W2PCLASGLVLPSEKLYHPCWPSVPLEVAPEFATSLVMLCIGESLVNSYNT 291
 0b 121 VHLWPGHAGSLVLPSEKLYHPCWPSVPLEVAPEFATSLVMLCIGESLVNSYNT 189
 0Y 272 PVTALGLLEDALEEFNTEFNINRVSSYSTASHRLRFGAGTIAELTTTAATPFMR 351
 0b 101 PVTALGLLEDALEEFNTEFNINRVSSYSTASHRLRFGAGTIAELTTTAATPFMR 249
 0Y 312 PLYPSTSPWVSTPRGTAALFNINATILGCLPTELLSSAGQPLSPRVVSSANGPTV 411
 0b 241 PLYPSTSPWVSTPRGTAALFNINATILGCLPTELLSSAGQPLSPRVVSSANGPTV 309
 0Y 474 LITAVERA:QWKAATHR:ELAG:SAVVLELQ:QH:ELKPTT:SG:AT:SP:SG:KARI:V: 471
 0b 601 LITAVERA:QWKAATHR:ELAG:SAVVLELQ:QH:ELKPTT:SG:AT:SP:SG:KARI:V: 369
 0Y 474 WLSH:TAAPV:AG:SYGSS:GPVYV:DSVTLVNVAG:AG:AVAS:KMT:KVI:LP:K:PLST:QQ 541
 0b 361 WLSH:TAAPV:AG:SYGSS:GPVYV:DSVTLVNVAG:AG:AVAS:KMT:KVI:LP:K:PLST:QQ 429
 0Y 542 TSE:TTV:V:LL:DEP:STW:AT:TEAT:TE:TH:HTA:PL:GL:V:V:NA:AG:HV:V:STTT:TC:AG 591

RESULT 15

AAR96103

1D AAR96103 standard; Protein; 52% AA.

XX AAR96103;

XX 05 AUG 1996 (first entry)

XX Hepatitis E virus (Burma strain) recombinant 62K antigen.

XX HEV: externally transmitted non-African hepatitis virus; vaccine;

XX diagnostic antigen; Spodoptera frugiperda; 51% insect;

XX baculovirus; capsid.

XX Hepatitis E virus Burma strain

XX W09612697 A2.

XX 02 MAY 1996.

XX 23 OCT 1995; 95W0514703.

XX 14 OCT 1995; 950S 0542644.

XX 24 OCT 1994; 940S 0427952.

XX (GENE) GENEJARS TECHNOLIGIES INC.

XX Fucst TR, McAtee CP, Yarbough PO, Zhang Y.

XX W01; 1996-230608/23.

XX Hepatitis E virus (HEV) antigens derived from opp 2 - useful as

XX diagnostic reagents for determining HEV infection and in vaccines

XX Claim 7; Page 95-97; 125pp; English.

XX DNA (AA127109) coding for the 62K antigen (AAR9601) of hepatitis E

XX virus (HEV) Burma strain capsid protein was cloned into baculovirus

XX expression vector pBaculatt and recombinant 62K was expressed

XX in Sf9 insect cells. High levels of expression were obtained and

XX the recombinant 62K was obtained in over 95% purity. However, C-terminal

XX processing resulted in the deletion of 9 or 23 amino acids from 62K,

XX giving 2 related polypeptide species (AAR96101 and AAR96103). Similar

XX results were obtained with HEV Mexico 62K antigen (see also AAR96102 and

XX AAR96104). Recombinant 62K represents an improved antigen, in

XX comparison to bacterial expressed proteins, for use in HEV

XX diagnostic assays, and also has excellent immunogenic properties.

XX Sequence 52% AA:

Query Match

Best Local Similarity 100.0%; Prod. No. 0;

Matches 52%; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

1 STATE: CA
2 COUNTRY: USA
3 ZIP: 94306
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: PatentIn Release #1.0, Version #1.2e
9
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US 07,977,941A
12 FILING DATE: 01-MAY-1992
13 CLASSIFICATION: 435
14
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US 822,335
17 FILING DATE: 17-JAN-1992
18
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: US 505,888
21 FILING DATE: 05-APRIL-1990
22
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 426,421
25 FILING DATE: 13-OCTOBER-1989
26
27 APPLICATION DATA:
28 APPLICATION NUMBER: US 367,486
29 FILING DATE: 16-JUNE-1989
30
31 APPLICATION DATA:
32 APPLICATION NUMBER: US 326,672
33 FILING DATE: 11-APRIL-1989
34
35 PRIOR APPLICATION DATA:
36 APPLICATION NUMBER: US 208,947
37 FILING DATE: 17-JUNE-1988
38
39 ATTORNEY/AGENT INFORMATION:
40 NAME: Sholluz, Charles K.
41 REGISTRATION NUMBER: 39,615
42 REFERENCE/DOCKET NUMBER: 4600-0093.33
43 TELECOMMUNICATION INFORMATION:
44 TELEPHONE: (415) 324-0880
45 TELEFAX: (415) 324-0960
46 INFORMATION FOR SEQ ID NO. 19:
47 SEQUENCE CHARACTERISTICS:
48 LENGTH: 660 amino acids
49 TYPE: amino acid
50 TOPOLOGY: linear
51 MOLECULE TYPE: protein
52 HYPOTHEICAL: NO
53 ORIGINAL SOURCE:
54 INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
55 US-07-876-941A-19
56
57 Query Match 100.0% Score 660; Hk 2; Length 660;
58 Best Local Similarity 100.0%; Pref No 0;
59 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
60
61 QY 1 MRPPIIIIIIMPIPMIPADPGQPS:PPPPPPSGSGGFWTQVLSAPPFAIPYIRTN 60
62 DE 1 MPPPIIIIIIMPIPMIPADPGQPS:PPPPPPSGSGGFWTQVLSAPPFAIPYIRTN 60
63 QY 61 PPAPVTAAGAGPPVPPVPPPIPGSAMPDQAGPPAVASPPPTTACAAFLTAAVADPNT 120
64 DE 61 PPAPVTAAGAGPPVPPVPPPIPGSAMPDQAGPPAVASPPPTTACAAFLTAAVADPNT 120
65 QY 121 PPVTAAGAGAGPPVPPVPPPIPGSAMPDQAGPPAVASPPPTTACAAFLTAAVADPNT 180
66 DE 121 PPVTAAGAGAGPPVPPVPPPIPGSAMPDQAGPPAVASPPPTTACAAFLTAAVADPNT 180
67 QY 181 NYAQYRVARATISRYPLVPMNAVGGYAIISIFWPTTTTPTSDMNSITSTDRHIIVQGI 240
68 DE 181 NYAQYRVARATISRYPLVPMNAVGGYAIISIFWPTTTTPTSDMNSITSTDRHIIVQGI 240
69 QY 241 ASPIVITISFPIHYPPNCPWSPVTSQVAFEATSGIVMCIHQSGIVNSYINTPYTGAI 300
70 DE 241 ASPIVITISFPIHYPPNCPWSPVTSQVAFEATSGIVMCIHQSGIVNSYINTPYTGAI 300
71 QY 301 DEPLEFRNLTPGNTNTPSPVSSSTAPRHPPLPAGAGCTAELITTAITPMPKITYETSTNG 360
72 DE 301 DEPLEFRNLTPGNTNTPSPVSSSTAPRHPPLPAGAGCTAELITTAITPMPKITYETSTNG 360
73
74 RESULT 5
75 US 08 479,537-H
76 Sequence 8, Application US/08479537
77 Patent No. 6120988
78 GENERAL INFORMATION:
79 APPLICANT: Reyes, Gregory R
80 APPLICANT: Yarbrough, Patricia O
81 APPLICANT: Bradley, Daniel W
82 ADDRESSEE: Frawood, Gregory J
83 APPLICANT: Tam, Albert
84 APPLICANT: Fitt, Kirk E
85 TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
86 TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
87 NUMBER OF SEQUENCES: 20
88 CORRESPONDENCE ADDRESS:
89 ADDRESSEE: Dohliador & Associates
90 STREET: 350 Cambridge Avenue, Suite 250
91 CITY: Palo Alto
92 STATE: CA
93 COUNTRY: USA
94 ZIP: 94306
95 MEDIUM TYPE: Floppy disk
96 COMPUTER: IBM PC compatible
97 SOFTWARE: PatentIn Release #1.0, Version #1.25
98 CURRENT APPLICATION DATA:
99 APPLICATION NUMBER: US 08,479,537
100 FILING DATE: 07-JUN-1995
101 PRIOR APPLICATION DATA:
102 APPLICATION NUMBER: US 08,479,537
103 FILING DATE: 25-JUL-1994
104 PRIOR APPLICATION DATA:
105 APPLICATION NUMBER: US 07,681,078
106 FILING DATE: 05-APR-1991
107 PRIOR APPLICATION DATA:
108 APPLICATION NUMBER: US 07,605,888
109 FILING DATE: 05-APR-1990
110 PRIOR APPLICATION DATA:
111 APPLICATION NUMBER: US 07,420,921
112 FILING DATE: 13-OCT-1989
113 PRIOR APPLICATION DATA:
114 APPLICATION NUMBER: US 07,403,486
115 FILING DATE: 16-JUN-1989
116 PRIOR APPLICATION DATA:
117 APPLICATION NUMBER: US 07,336,632
118 FILING DATE: 11-APR-1989
119 PRIOR APPLICATION DATA:
120 APPLICATION NUMBER: US 07,208,947

```



```

1  SEQUENCE SYSTEM: 100 IPS/MS 100S
2  CURRENT APPLICATION DATA:
3  APPLICATION NUMBER: 05/07/7675, 08CA
4  FILING DATE: 20 APRIL 1992
5  CLASSIFICATION: 435
6  PRIOR APPLICATION DATA:
7  APPLICATION NUMBER: US 0422,435
8  FILING DATE: 17 JAN 1992
9  FILING DATE: 15 APRIL 1999
10 APPLICATION NUMBER: US 506,898
11 FILING DATE: 05 APRIL 1999
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: US 428,921
14 FILING DATE: 15 OCTOBER 1989
15 APPLICATION DATA:
16 APPLICATION NUMBER: US 467,486
17 FILING DATE: 16 JUNE 1989
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 439,672
20 FILING DATE: 11 APRIL 1989
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 208,997
23 FILING DATE: 17 JUNE 1988
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Fabiano, Gary R.
26 REGISTRATION NUMBER: 43,875
27 REFERENCE/WORKSHEET NUMBER: 4600 0195, 430
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (415) 424-0980
30 TELEFAX: (415) 424-0960
31 INFORMATION FOR SEQ ID NO: 19:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 600 amino acids
34 TYPE: amino acid
35 MOLECULE TYPE: protein
36 HYDROPHILIC: N
37 ORIGINAL SOURCE:
38 INDIVIDUAL ISOLATE: 08F 2, BORMA, FIGURE 9
39 US 07 851-410A 19

Query Match: 100.00% Score: 660; ID: 4; Length: 600;
Best Local Similarity: 100.00% Ident. No.: 0;
Matches: 660; Conservative: 0; Mismatches: 0; Gaps: 0;

27 1 MRPPPTLLMLPMPAPAPGQPSGRGGRGSGSGGFWCDRWISQPPALPYTHPTN 60
28 1 MRPPPTLLMLPMPAPAPGQPSGRGGRGSGSGGFWCDRWISQPPALPYTHPTN 60
29 61 PFAPVIAAAGAGRVQAPAPGASAMPQQAQPPAVASRRPTTAGAAPLTAVAPADTP 120
30 61 PFAPVIAAAGAGRVQAPAPGASAMPQQAQPPAVASRRPTTAGAAPLTAVAPADTP 120
31 121 FVPIVASGALLRPQYNLSISLSSVAIGTNILVYAAULSPILPQDGTNTHMATEAS 180
32 121 FVPIVASGALLRPQYNLSISLSSVAIGTNILVYAAULSPILPQDGTNTHMATEAS 180
33 121 FVPIVASGALLRPQYNLSISLSSVAIGTNILVYAAULSPILPQDGTNTHMATEAS 180
34 121 FVPIVASGALLRPQYNLSISLSSVAIGTNILVYAAULSPILPQDGTNTHMATEAS 180
35 121 FVPIVASGALLRPQYNLSISLSSVAIGTNILVYAAULSPILPQDGTNTHMATEAS 180
36 121 FVPIVASGALLRPQYNLSISLSSVAIGTNILVYAAULSPILPQDGTNTHMATEAS 180
37 241 ASSELVTSERPHVQNGVRSVETSGVAEEATSGVLMGCTHGSLSVSYNTPTTGALL 400
38 241 ASSELVTSERPHVQNGVRSVETSGVAEEATSGVLMGCTHGSLSVSYNTPTTGALL 400
39 400 400
40 400 400
41 400 400
42 400 400
43 400 400
44 400 400
45 400 400
46 400 400
47 400 400
48 400 400
49 400 400
50 400 400
51 400 400
52 400 400
53 400 400
54 400 400
55 400 400
56 400 400
57 400 400
58 400 400
59 400 400
60 400 400
61 400 400
62 400 400
63 400 400
64 400 400
65 400 400
66 400 400
67 400 400
68 400 400
69 400 400
70 400 400
71 400 400
72 400 400
73 400 400
74 400 400
75 400 400
76 400 400
77 400 400
78 400 400
79 400 400
80 400 400
81 400 400
82 400 400
83 400 400
84 400 400
85 400 400
86 400 400
87 400 400
88 400 400
89 400 400
90 400 400
91 400 400
92 400 400
93 400 400
94 400 400
95 400 400
96 400 400
97 400 400
98 400 400
99 400 400
100 400 400
101 400 400
102 400 400
103 400 400
104 400 400
105 400 400
106 400 400
107 400 400
108 400 400
109 400 400
110 400 400
111 400 400
112 400 400
113 400 400
114 400 400
115 400 400
116 400 400
117 400 400
118 400 400
119 400 400
120 400 400
121 400 400
122 400 400
123 400 400
124 400 400
125 400 400
126 400 400
127 400 400
128 400 400
129 400 400
130 400 400
131 400 400
132 400 400
133 400 400
134 400 400
135 400 400
136 400 400
137 400 400
138 400 400
139 400 400
140 400 400
141 400 400
142 400 400
143 400 400
144 400 400
145 400 400
146 400 400
147 400 400
148 400 400
149 400 400
150 400 400
151 400 400
152 400 400
153 400 400
154 400 400
155 400 400
156 400 400
157 400 400
158 400 400
159 400 400
160 400 400
161 400 400
162 400 400
163 400 400
164 400 400
165 400 400
166 400 400
167 400 400
168 400 400
169 400 400
170 400 400
171 400 400
172 400 400
173 400 400
174 400 400
175 400 400
176 400 400
177 400 400
178 400 400
179 400 400
180 400 400
181 400 400
182 400 400
183 400 400
184 400 400
185 400 400
186 400 400
187 400 400
188 400 400
189 400 400
190 400 400
191 400 400
192 400 400
193 400 400
194 400 400
195 400 400
196 400 400
197 400 400
198 400 400
199 400 400
200 400 400
201 400 400
202 400 400
203 400 400
204 400 400
205 400 400
206 400 400
207 400 400
208 400 400
209 400 400
210 400 400
211 400 400
212 400 400
213 400 400
214 400 400
215 400 400
216 400 400
217 400 400
218 400 400
219 400 400
220 400 400
221 400 400
222 400 400
223 400 400
224 400 400
225 400 400
226 400 400
227 400 400
228 400 400
229 400 400
230 400 400
231 400 400
232 400 400
233 400 400
234 400 400
235 400 400
236 400 400
237 400 400
238 400 400
239 400 400
240 400 400
241 400 400
242 400 400
243 400 400
244 400 400
245 400 400
246 400 400
247 400 400
248 400 400
249 400 400
250 400 400
251 400 400
252 400 400
253 400 400
254 400 400
255 400 400
256 400 400
257 400 400
258 400 400
259 400 400
260 400 400
261 400 400
262 400 400
263 400 400
264 400 400
265 400 400
266 400 400
267 400 400
268 400 400
269 400 400
270 400 400
271 400 400
272 400 400
273 400 400
274 400 400
275 400 400
276 400 400
277 400 400
278 400 400
279 400 400
280 400 400
281 400 400
282 400 400
283 400 400
284 400 400
285 400 400
286 400 400
287 400 400
288 400 400
289 400 400
290 400 400
291 400 400
292 400 400
293 400 400
294 400 400
295 400 400
296 400 400
297 400 400
298 400 400
299 400 400
300 400 400
301 400 400
302 400 400
303 400 400
304 400 400
305 400 400
306 400 400
307 400 400
308 400 400
309 400 400
310 400 400
311 400 400
312 400 400
313 400 400
314 400 400
315 400 400
316 400 400
317 400 400
318 400 400
319 400 400
320 400 400
321 400 400
322 400 400
323 400 400
324 400 400
325 400 400
326 400 400
327 400 400
328 400 400
329 400 400
330 400 400
331 400 400
332 400 400
333 400 400
334 400 400
335 400 400
336 400 400
337 400 400
338 400 400
339 400 400
340 400 400
341 400 400
342 400 400
343 400 400
344 400 400
345 400 400
346 400 400
347 400 400
348 400 400
349 400 400
350 400 400
351 400 400
352 400 400
353 400 400
354 400 400
355 400 400
356 400 400
357 400 400
358 400 400
359 400 400
360 400 400
361 400 400
362 400 400
363 400 400
364 400 400
365 400 400
366 400 400
367 400 400
368 400 400
369 400 400
370 400 400
371 400 400
372 400 400
373 400 400
374 400 400
375 400 400
376 400 400
377 400 400
378 400 400
379 400 400
380 400 400
381 400 400
382 400 400
383 400 400
384 400 400
385 400 400
386 400 400
387 400 400
388 400 400
389 400 400
390 400 400
391 400 400
392 400 400
393 400 400
394 400 400
395 400 400
396 400 400
397 400 400
398 400 400
399 400 400
400 400 400
401 400 400
402 400 400
403 400 400
404 400 400
405 400 400
406 400 400
407 400 400
408 400 400
409 400 400
410 400 400
411 400 400
412 400 400
413 400 400
414 400 400
415 400 400
416 400 400
417 400 400
418 400 400
419 400 400
420 400 400
421 400 400
422 400 400
423 400 400
424 400 400
425 400 400
426 400 400
427 400 400
428 400 400
429 400 400
430 400 400
431 400 400
432 400 400
433 400 400
434 400 400
435 400 400
436 400 400
437 400 400
438 400 400
439 400 400
440 400 400
441 400 400
442 400 400
443 400 400
444 400 400
445 400 400
446 400 400
447 400 400
448 400 400
449 400 400
450 400 400
451 400 400
452 400 400
453 400 400
454 400 400
455 400 400
456 400 400
457 400 400
458 400 400
459 400 400
460 400 400
461 400 400
462 400 400
463 400 400
464 400 400
465 400 400
466 400 400
467 400 400
468 400 400
469 400 400
470 400 400
471 400 400
472 400 400
473 400 400
474 400 400
475 400 400
476 400 400
477 400 400
478 400 400
479 400 400
480 400 400
481 400 400
482 400 400
483 400 400
484 400 400
485 400 400
486 400 400
487 400 400
488 400 400
489 400 400
490 400 400
491 400 400
492 400 400
493 400 400
494 400 400
495 400 400
496 400 400
497 400 400
498 400 400
499 400 400
500 400 400
501 400 400
502 400 400
503 400 400
504 400 400
505 400 400
506 400 400
507 400 400
508 400 400
509 400 400
510 400 400
511 400 400
512 400 400
513 400 400
514 400 400
515 400 400
516 400 400
517 400 400
518 400 400
519 400 400
520 400 400
521 400 400
522 400 400
523 400 400
524 400 400
525 400 400
526 400 400
527 400 400
528 400 400
529 400 400
530 400 400
531 400 400
532 400 400
533 400 400
534 400 400
535 400 400
536 400 400
537 400 400
538 400 400
539 400 400
540 400 400
541 400 400
542 400 400
543 400 400
544 400 400
545 400 400
546 400 400
547 400 400
548 400 400
549 400 400
550 400 400
551 400 400
552 400 400
553 400 400
554 400 400
555 400 400
556 400 400
557 400 400
558 400 400
559 400 400
560 400 400
561 400 400
562 400 400
563 400 400
564 400 400
565 400 400
566 400 400
567 400 400
568 400 400
569 400 400
570 400 400
571 400 400
572 400 400
573 400 400
574 400 400
575 400 400
576 400 400
577 400 400
578 400 400
579 400 400
580 400 400
581 400 400
582 400 400
583 400 400
584 400 400
585 400 400
586 400 400
587 400 400
588 400 400
589 400 400
590 400 400
591 400 400
592 400 400
593 400 400
594 400 400
595 400 400
596 400 400
597 400 400
598 400 400
599 400 400
600 400 400
601 400 400
602 400 400
603 400 400
604 400 400
605 400 400
606 400 400
607 400 400
608 400 400
609 400 400
610 400 400
611 400 400
612 400 400
613 400 400
614 400 400
615 400 400
616 400 400
617 400 400
618 400 400
619 400 400
620 400 400
621 400 400
622 400 400
623 400 400
624 400 400
625 400 400
626 400 400
627 400 400
628 400 400
629 400 400
630 400 400
631 400 400
632 400 400
633 400 400
634 400 400
635 400 400
636 400 400
637 400 400
638 400 400
639 400 400
640 400 400
641 400 400
642 400 400
643 400 400
644 400 400
645 400 400
646 400 400
647 400 400
648 400 400
649 400 400
650 400 400
651 400 400
652 400 400
653 400 400
654 400 400
655 400 400
656 400 400
657 400 400
658 400 400
659 400 400
660 400 400
661 400 400
662 400 400
663 400 400
664 400 400
665 400 400
666 400 400
667 400 400
668 400 400
669 400 400
670 400 400
671 400 400
672 400 400
673 400 400
674 400 400
675 400 400
676 400 400
677 400 400
678 400 400
679 400 400
680 400 400
681 400 400
682 400 400
683 400 400
684 400 400
685 400 400
686 400 400
687 400 400
688 400 400
689 400 400
690 400 400
691 400 400
692 400 400
693 400 400
694 400 400
695 400 400
696 400 400
697 400 400
698 400 400
699 400 400
700 400 400
701 400 400
702 400 400
703 400 400
704 400 400
705 400 400
706 400 400
707 400 400
708 400 400
709 400 400
710 400 400
711 400 400
712 400 400
713 400 400
714 400 400
715 400 400
716 400 400
717 400 400
718 400 400
719 400 400
720 400 400
721 400 400
722 400 400
723 400 400
724 400 400
725 400 400
726 400 400
727 400 400
728 400 400
729 400 400
730 400 400
731 400 400
732 400 400
733 400 400
734 400 400
735 400 400
736 400 400
737 400 400
738 400 400
739 400 400
740 400 400
741 400 400
742 400 400
743 400 400
744 400 400
745 400 400
746 400 400
747 400 400
748 400 400
749 400 400
750 400 400
751 400 400
752 400 400
753 400 400
754 400 400
755 400 400
756 400 400
757 400 400
758 400 400
759 400 400
760 400 400
761 400 400
762 400 400
763 400 400
764 400 400
765 400 400
766 400 400
767 400 400
768 400 400
769 400 400
770 400 400
771 400 400
772 400 400
773 400 400
774 400 400
775 400 400
776 400 400
777 400 400
778 400 400
779 400 400
780 400 400
781 400 400
782 400 400
783 400 400
784 400 400
785 400 400
786 400 400
787 400 400
788 400 400
789 400 400
790 400 400
791 400 400
792 400 400
793 400 400
794 400 400
795 400 400
796 400 400
797 400 400
798 400 400
799 400 400
800 400 400
801 400 400
802 400 400
803 400 400
804 400 400
805 400 400
806 400 400
807 400 400
808 400 400
809 400 400
810 400 400
811 400 400
812 400 400
813 400 400
814 400 400
815 400 400
816 400 400
817 400 400
818 400 400
819 400 400
820 400 400
821 400 400
822 400 400
823 400 400
824 400 400
825 400 400
826 400 400
827 400 400
828 400 400
829 400 400
830 400 400
831 400 400
832 400 400
833 400 400
834 400 400
835 400 400
836 400 400
837 400 400
838 400 400
839 400 400
840 400 400
841 400 400
842 400 400
843 400 400
844 400 400
845 400 400
846 400 400
847 400 400
848 400 400
849 400 400
850 400 400
851 400 400
852 400 400
853 400 400
854 400 400
855 400 400
856 400 400
857 400 400
858 400 400
859 400 400
860 400 400
861 400 400
862 400 400
863 400 400
864 400 400
865 400 400
866 400 400
867 400 400
868 400 400
869 400 400
870 400 400
871 400 400
872 400 400
873 400 400
874 400 400
875 400 400
876 400 400
877 400 400
878 400 400
879 400 400
880 400 400
881 400 400
882 400 400
883 400 400
884 400 400
885 400 400
886 400 400
887 400 400
888 400 400
889 400 400
890 400 400
891 400 400
892 400 400
893 400 400
894 400 400
895 400 400
896 400 400
897 400 400
898 400 400
899 400 400
900 400 400
901 400 400
902 400 400
903 400 400
904 400 400
905 400 400
906 400 400
907 400 400
908 400 400
909 400 400
910 400 400
911 400 400
912 400 400
913 400 400
914 400 400
915 400 400
916 400 400
917 400 400
918 400 400
919 400 400
920 400 400
921 400 400
922 400 400
923 400 400
924 400 400
925 400 400
926 400 400
927 400 400
928 400 400
929 400 400
930 400 400
931 400 400
932 400 400
933 400 400
934 400 400
935 400 400
936 400 400
937 400 400
938 400 400
939 400 400
940 400 400
941 400 400
942 400 400
943 400 400
944 400 400
945 400 400
946 400 400
947 400 400
948 400 400
949 400 400
950 400 400
951 400 400
952 400 400
953 400 400
954 400 400
955 400 400
956 400 400
957 400 400
958 400 400
959 400 400
960 400 400
961 400 400
962 400 400
963 400 400
964 400 400
965 400 400
966 400 400
967 400 400
968 400 400
969 4
```

121 PVPDVSRCATIRREGYMLSTSPSSVATGNTLVYAAELSPLLPQRTNTHIMATEAS 180
121 PVPDVSRCALLRRQNLSTSPSSVATGNTLVYAAELSPLLPQRTNTHIMATEAS 180
181 NYAQYRVARATIRYRLVFNAGGYAISISFWPQTITPTISVIMNSITSHVPLVQPEI 240
181 NYAQYRVARATIRYRLVFNAGGYAISISFWPQTITPTISVIMNSITSHVPLVQPEI 240
241 ASELVPSERLHYRNQGRSVETSGVAFREATSGLVMLCIGHSLVNSVNTPTTCALGIL 300
241 ASELVPSERLHYRNQGRSVETSGVAFREATSGLVMLCIGHSLVNSVNTPTTCALGIL 300
301 DFALELEFERNLTPONTNTRVSRYSSTAPHRIPGAGTAFELTTTAAATPFMKLYPTSTNG 360
301 DFALELEFERNLTPONTNTRVSRYSSTAPHRIPGAGTAFELTTTAAATPFMKLYPTSTNG 360
361 VRTGRTATITINLADTLGRLPTPISSAGSQI PPSPPVVSANGPTVKLYTSVENAQ 420
361 VRTGRTATITINLADTLGRLPTPISSAGSQI PPSPPVVSANGPTVKLYTSVENAQ 420
421 QDKGIALPHIDIGESKVVQDDVNDQHEGERTPSAPSPKPSVLKANDVLMLSITAAFY 480
421 QDKGIALPHIDIGESKVVQDDVNDQHEGERTPSAPSPKPSVLKANDVLMLSITAAFY 480
481 DQSTYSSSTGPVYVSDSVTLNVNATCAQAVAPSLWTKVTLGDPPLSTIQYVSKTFPVL 540
481 DQSTYSSSTGPVYVSDSVTLNVNATCAQAVAPSLWTKVTLGDPPLSTIQYVSKTFPVL 540
541 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
541 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
601 LAPSALALLIEDILYAPARHTPDGPPPTPRGLGQAFQPSIVAFQPKMKVCKRREL 660
601 LAPSALALLIEDILYAPARHTPDGPPPTPRGLGQAFQPSIVAFQPKMKVCKRREL 660

RESULT 14
US-08-542-644-15
: Sequence 15, Application US-08-542-644
: Patent No. 6214970
: GENERAL INFORMATION:
: APPLICANT: Fuerst, Thomas R.
: APPLICANT: McAtee, C. Patrick
: APPLICANT: Yarbough, Patricia O.
: APPLICANT: Zhang, Yifan
: TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave., Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: ProDOS/Me-Pos
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US-08-542-644
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Fabian, Gary R.
: REGISTRATION NUMBER: 33,875
: REFERENCE/BOOKLET NUMBER: 4600-0203-20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 324-0880
: TELEFAX: (415) 324-0960
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:

LENGTH: 542 amino acids
: TYPE: amino acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: protein
: HYDROPHILIC: No
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
: INDIVIDUAL ISOLATE: 16.8.1, FIGURE 4
US-08-542-644-15
Query Match 84.2% Score 549; DB 3; Length 549;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 112 AVAPADHTVPPVNVISRGALRRQNLSTSPSSVATGNTLVYAAELSPLLPQRTNTHIMATEAS 171
Db 1 AVAPADHTVPPVNVISRGALRRQNLSTSPSSVATGNTLVYAAELSPLLPQRTNTHIMATEAS 171
QY 172 THIMATASNYAQYRVARATIRYRLVFNAGGYAISISFWPQTITPTISVIMNSITSHVPLVQPEI 241
Db 61 THIMATASNYAQYRVARATIRYRLVFNAGGYAISISFWPQTITPTISVIMNSITSHVPLVQPEI 241
QY 242 VRTGRTATITINLADTLGRLPTPISSAGSQI PPSPPVVSANGPTVKLYTSVENAQ 420
Db 121 VRTGRTATITINLADTLGRLPTPISSAGSQI PPSPPVVSANGPTVKLYTSVENAQ 420
QY 292 FYICAGLLIDFALEFERNLTPONTNTRVSRYSSTAPHRIPGAGTAFELTTTAAATPFMKLYPTSTNG 361
Db 181 FYICAGLLIDFALEFERNLTPONTNTRVSRYSSTAPHRIPGAGTAFELTTTAAATPFMKLYPTSTNG 361
QY 411 DQSTYSSSTGPVYVSDSVTLNVNATCAQAVAPSLWTKVTLGDPPLSTIQYVSKTFPVL 540
Db 241 DQSTYSSSTGPVYVSDSVTLNVNATCAQAVAPSLWTKVTLGDPPLSTIQYVSKTFPVL 540
QY 471 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
Db 301 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
QY 541 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
Db 361 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
QY 591 VRTGRTATITINLADTLGRLPTPISSAGSQI PPSPPVVSANGPTVKLYTSVENAQ 420
Db 421 VRTGRTATITINLADTLGRLPTPISSAGSQI PPSPPVVSANGPTVKLYTSVENAQ 420
QY 651 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
Db 481 LKCKLSFWACTTKAGYPYNYNTASQILLVNAACHRVASTYTTTSIGAGPVSISAVAV 600
QY 652 MKVCKRREL 660
Db 541 MKVCKRREL 549

RESULT 14
US-08-477-292-15
: Sequence 15, Application US-08-477-292
: Patent No. 6214641
: GENERAL INFORMATION:
: APPLICANT: Fuerst, Thomas
: APPLICANT: McAtee, Patrick
: APPLICANT: Yarbough, Patricia
: APPLICANT: Zhang, Yifan
: TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND METHODS
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kathleen M. Desjardins, M.D.
: STREET: 505 Penobscot Drive
: CITY: Redwood City
: STATE: CA
: COUNTRY: US

1 ZIPS: 4000
 2 COMPUTER READABLE FORM:
 3 MEDIUM TYPE: floppy disk
 4 COMPUTER: IBM pc compatible
 5 OPERATING SYSTEM: PC-DOS/MS-DOS
 6 SOFTWARE: Patent Release #1.0, Version #1.25
 7 CURRENT APPLICATION DATA:
 8 APPLICATION NUMBER: US598477, 292
 9 FILING DATE:
 10 CLASSIFICATION:
 11 PRIOR APPLICATION DATA:
 12 APPLICATION NUMBER: 05/08/427, 952
 13 FILING DATE:
 14 ATTORNEY/AGENT INFORMATION:
 15 NAME: Brooks, Allan A.
 16 REGISTRATION NUMBER: 46, 074
 17 FEE REFERENCE NUMBER: 5215
 18 TELECOMMUNICATION INFORMATION:
 19 TELEPHONE: (415) 699, 9500
 20 TELEFAX: (415) 668, 0709
 21 INFORMATION FOR SEQ ID NO: 15:
 22 LENGTH: 549 amino acids
 23 TYPE: amino acid
 24 STRANDEDNESS: single
 25 TOPOLOGY: linear
 26 MOLECULE TYPE: protein
 27 HYDROPHILIC: No
 28 ORIGINAL SOURCE:
 29 INDIVIDUAL ISOLATE: Hepatitis E Virus (Hepatitis E Virus)
 30 INDIVIDUAL ISOLATE: F62kda, FIGURE 4
 31 US 08 477 292 15

Query Match: 84.28; Score 549; Ldb 4; Length 549;
 Best Local Similarity: 100.00; Pred. No. 0;
 Matches 549; Conservat 0; Missed 0; Models 0; Gaps 0

07 112 AVAABDPPFVWDSRCALIRKQYNLSTSLSSVATGIVLYAAPIPLPLOGIN 171
 10 1 AVAABDPPFVWDSRCALIRKQYNLSTSLSSVATGIVLYAAPIPLPLOGIN 60

07 172 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 241
 10 1 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 60

07 61 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 120
 10 1 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 60

07 272 VLLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 251
 10 1 VLLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 60

07 121 VPLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 180
 10 1 VPLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 60

07 272 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 451
 10 1 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 60

07 121 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 240
 10 1 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 60

07 572 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 411
 10 1 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 60

07 241 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 300
 10 1 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 60

07 412 LYSVVENACQFTLALPRLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 471
 10 1 LYSVVENACQFTLALPRLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 60

07 472 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 541
 10 1 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 60

07 541 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 420
 10 1 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 60

07 572 YSTVFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 591
 10 1 YSTVFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 60

07 421 YSKTFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 480
 10 1 YSKTFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 60

07 572 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 651
 10 1 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 60

07 401 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 540
 10 1 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 60

07 652 MKVQKREPL 660
 10 1 MKVQKREPL 540

RESULT 15
 PCT US95 14704 15
 1 Sequence 15, Application 15/105953, 15
 2 GENERAL INFORMATION:
 3 APPLICANT:
 4 TITLE OF INVENTION: IDENTIFICATION OF VIRUS ANTIGEN AND
 5 TITLE OF INVENTION: USES THEREOF
 6 NUMBER OF SEQUENCES: 41
 7 CORRESPONDENT ADDRESS:
 8 ADDRESSEE: Schindler & Associates
 9 STREET: P.O. Box 60860
 10 CITY: Palo Alto
 11 STATE: CA
 12 COUNTRY: USA
 13 ZIP: 94306, 060
 14 COMPUTER READABLE FORM:
 15 MEDIUM TYPE: floppy disk
 16 COMPUTER: IBM pc compatible
 17 OPERATING SYSTEM: PC-DOS/MS-DOS
 18 SOFTWARE: Patent Release #1.0, Version #1.25
 19 CURRENT APPLICATION DATA:
 20 APPLICATION NUMBER: 05/08/477, 292
 21 FILING DATE:
 22 CLASSIFICATION:
 23 PRIOR APPLICATION DATA:
 24 APPLICATION NUMBER: 05/08/427, 952
 25 FILING DATE:
 26 ATTORNEY/AGENT INFORMATION:
 27 NAME: Brooks, Allan A.
 28 REGISTRATION NUMBER: 46, 074
 29 FEE REFERENCE NUMBER: 5215
 30 TELECOMMUNICATION INFORMATION:
 31 TELEPHONE: (415) 699, 9500
 32 TELEFAX: (415) 668, 0709
 33 INFORMATION FOR SEQ ID NO: 15:
 34 LENGTH: 549 amino acids
 35 TYPE: amino acid
 36 STRANDEDNESS: single
 37 TOPOLOGY: linear
 38 MOLECULE TYPE: protein
 39 HYDROPHILIC: No
 40 ORIGINAL SOURCE:
 41 INDIVIDUAL ISOLATE: Hepatitis E Virus (Hepatitis E Virus)
 42 INDIVIDUAL ISOLATE: F62kda, FIGURE 4
 43 PCT US95 14704 15

Query Match: 84.28; Score 549; Ldb 4; Length 549;
 Best Local Similarity: 100.00; Pred. No. 0;
 Matches 549; Conservat 0; Missed 0; Models 0; Gaps 0

07 112 AVAABDPPFVWDSRCALIRKQYNLSTSLSSVATGIVLYAAPIPLPLOGIN 171
 10 1 AVAABDPPFVWDSRCALIRKQYNLSTSLSSVATGIVLYAAPIPLPLOGIN 60

07 172 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 241
 10 1 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 60

07 61 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 120
 10 1 THMALASNAQYKRAKATRYPLVNAVGYAISTSWQFTTPTSVIMNSTSTD 60

07 272 VLLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 251
 10 1 VLLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 60

07 121 VPLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 180
 10 1 VPLVQVQACAEVTPFERHRRQKSWETGVVAHREATSLVYVLCRGGSLVRYRT 60

07 272 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 451
 10 1 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 60

07 121 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 240
 10 1 PYGADGILDFALDEPFRNLPGNTNTRVSPYSTAHPLRRCAGDCTAELTTAAATPKM 60

07 572 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 411
 10 1 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 60

07 241 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 300
 10 1 GLYVLSINQWELTRGALLLPLRLADLLGLLTELISAAAGGLFYSRWVVSANGELPVK 60

07 412 LYSVVENACQFTLALPRLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 471
 10 1 LYSVVENACQFTLALPRLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 60

07 472 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 541
 10 1 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 60

07 541 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 420
 10 1 WLSLLAARYDGLSTWSTGSPVYVSVLWVNAVALAAVAPSLTPVFTDQDPSITDQ 60

07 572 YSTVFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 591
 10 1 YSTVFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 60

07 421 YSKTFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 480
 10 1 YSKTFTVPLDGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLLGLL 60

07 572 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 651
 10 1 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 60

07 401 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 540
 10 1 PVSLSAVAVLHSAALLEDLGLYFAKATHTDGLFGRGLGLQGVAFQSLVAVLQRLK 60

412	LYTVFNQAQDKRTIAPHDDTLQPSRVLIQVYNQHEQDQPPSPAPSPFPFSVI	PANIWI	471
413			
414			
415			
416			
417	LYTSVENAQQDKRTIAPHDDTLQPSRVLIQVYNQHEQDQPPSPAPSPFPFSVL	PANDVL	360
418			
419			
420			
421	WLSLTAAYDQSTGSGSTGPGVYSDSVLIWNATAGAVARSLDWTKVTL	DGRPLSTTQQ	531
422			
423			
424			
425	WLSLTAAYDQSTGSGSTGPGVYSDSVLIWNATAGAVARSLDWTKVTL	DGRPLSTTQQ	420
426			
427	YKSTFFVLPLRGKLSFWAGCTTKAGYPYNYNTTASDQLLVENAGHRAVAIS	TYTTSLGAG	591
428			
429			
430			
431	YKSTFFVLPLRGKLSFWAGCTTKAGYPYNYNTTASDQLLVENAGHRAVAIS	TYTTSLGAG	480
432			
433			
434	PVLSISAVAVLAPHASALALLEDITLDYPARAHTFDHCPGCPPIGLIGGCA	FQSTVAELQPLK	651
435			
436			
437	PVLSISAVAVLAPHASALALLEDITLDYPARAHTFDHCPGCPPIGLIGGCA	FQSTVAELQMLK	540
438			
439			
440			
441			
442			
443			
444			
445			
446			
447			
448			
449			
450			
451			
452			
453			
454			
455			
456			
457			
458			
459			
460			
461			
462			
463			
464			
465			
466			
467			
468			
469			
470			
471			
472			
473			
474			
475			
476			
477			
478			
479			
480			
481			
482			
483			
484			
485			
486			
487			
488			
489			
490			
491			
492			
493			
494			
495			
496			
497			
498			
499			
500			
501			
502			
503			
504			
505			
506			
507			
508			
509			
510			
511			
512			
513			
514			
515			

Search completed: August 5, 2003, 09:34:36
Job time : 20 secs

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/542,644
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/ACKPT NUMBR: 4600-02943 30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 540 amino acids
 TYPE: amino acid
 STRANDEDNESS: Hepatitis E virus (Burma strain)
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:
 US-09-769-966-25

Query Match 81.8% Score 540; DB 10; Length 540;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 AVAPAHPTPPVPIVDSRGALLRQYNLSTSPITSSVATGTLNVLVYAAPLPLPLQDSTN 171
 Db 1 AVAPAHPTPPVPIVDSRGALLRQYNLSTSPITSSVATGTLNVLVYAAPLPLPLQDSTN 60
 QY 172 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTITTPISVMNISTD 231
 Db 61 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTITTPISVMNISTD 120
 QY 232 VRILVQGIASELIVISEKLVHPCQWPSVETSGVAEEATSGLVMLC IHGSLVNSYNT 291
 Db 121 VRILVQGIASELIVISEKLVHPCQWPSVETSGVAEEATSGLVMLC IHGSLVNSYNT 180
 QY 292 PYTGALGIDFALIFEPNITPNTNTVPSYSSSTAPRPLPGAPGIAFLITTAATPMK 351
 Db 181 PYTGALGIDFALIFEPNITPNTNTVPSYSSSTAPRPLPGAPGIAFLITTAATPMK 240
 QY 352 DLYPTSTNGVCFIQCIGIALITLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVR 411
 Db 241 DLYPTSTNGVCFIQCIGIALITLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVR 360
 QY 411 LYSVFNACQDQKGIAPRDTIGESPVVITQYDNGHEGLRPTSPAPSPSPVIANVI 471
 Db 360 LYSVFNACQDQKGIAPRDTIGESPVVITQYDNGHEGLRPTSPAPSPSPVIANVI 320
 QY 472 WLSLTAAPVQSTYGSSTPPVVSUSVITVNVATCAQAVPSI FWTPTVTLDPEN STIQQ 531
 Db 321 WLSLTAAPVQSTYGSSTPPVVSUSVITVNVATCAQAVPSI FWTPTVTLDPEN STIQQ 420
 QY 532 YKSTFWV LSPCVI SPWIACTKACQVYVYNTASQI IVERACIPVAISTYTSIGAG 591
 Db 420 YKSTFWV LSPCVI SPWIACTKACQVYVYNTASQI IVERACIPVAISTYTSIGAG 480
 QY 592 PVSISAVAVLAPHSALAL ETLIDVYPAKATFDQDQDNERDGLGCGAFQSTVAELKRL 651
 Db 480 PVSISAVAVLAPHSALAL ETLIDVYPAKATFDQDQDNERDGLGCGAFQSTVAELKRL 640

RESULT 5
 US-09-769-966-27
 Sequence 27: Application US/08/542,644
 Patent RE US/08/542,644
 GENERAL INFORMATION:
 APPLICANT: Fuerst, Thomas P.
 McAtee, C. Patrick
 Yarbough, Patricia G.
 Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 350 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC BUS/MS DOS
 SOFTWARE: Patent In Polyase #1.0, Version #1.25
 CURRENT APPLICATION DATA: 08/542,644
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/ACKPT NUMBR: 4600-02943 30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 525 amino acids
 TYPE: amino acid
 STRANDEDNESS: Hepatitis E virus (Burma strain)
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 SEQUENCE DESCRIPTION: SEQ ID NO: 27:
 US-09-769-966-27

Query Match 79.5% Score 525; DB 10; Length 525;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 AVAARAGGVIVSMSSGALHFEATRIEITELSEVATTRVVAARAPNLPVLSSTN 171
 Db 1 AVAARAGGVIVSMSSGALHFEATRIEITELSEVATTRVVAARAPNLPVLSSTN 60
 QY 172 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTITTPISVMNISTD 231
 Db 61 THIMATEASNYAQRVARATIRYRPLVNAVGGYATISFWPQTITTPISVMNISTD 120
 QY 232 VRILVQGIASELIVISEKLVHPCQWPSVETSGVAEEATSGLVMLC IHGSLVNSYNT 291
 Db 121 VRILVQGIASELIVISEKLVHPCQWPSVETSGVAEEATSGLVMLC IHGSLVNSYNT 180
 QY 292 PYTGALGIDFALIFEPNITPNTNTVPSYSSSTAPRPLPGAPGIAFLITTAATPMK 351
 Db 181 PYTGALGIDFALIFEPNITPNTNTVPSYSSSTAPRPLPGAPGIAFLITTAATPMK 240
 QY 352 DLYPTSTNGVCFIQCIGIALITLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVR 411
 Db 241 DLYPTSTNGVCFIQCIGIALITLNLADTLIGLPTLISAGQGLPSPPVVSANGPTVR 360
 QY 411 LYSVFNACQDQKGIAPRDTIGESPVVITQYDNGHEGLRPTSPAPSPSPVIANVI 471
 Db 360 LYSVFNACQDQKGIAPRDTIGESPVVITQYDNGHEGLRPTSPAPSPSPVIANVI 320
 QY 472 WLSLTAAPVQSTYGSSTPPVVSUSVITVNVATCAQAVPSI FWTPTVTLDPEN STIQQ 531
 Db 321 WLSLTAAPVQSTYGSSTPPVVSUSVITVNVATCAQAVPSI FWTPTVTLDPEN STIQQ 420
 QY 532 YKSTFWV LSPCVI SPWIACTKACQVYVYNTASQI IVERACIPVAISTYTSIGAG 591
 Db 420 YKSTFWV LSPCVI SPWIACTKACQVYVYNTASQI IVERACIPVAISTYTSIGAG 480
 QY 592 PVSISAVAVLAPHSALAL ETLIDVYPAKATFDQDQDNERDGLGCGAFQSTVAELKRL 651
 Db 480 PVSISAVAVLAPHSALAL ETLIDVYPAKATFDQDQDNERDGLGCGAFQSTVAELKRL 640

QY 224 MNSITSTVDRIIVQGIASELVIPSERHYRNGQWRSVETSGVAEEATSLVWMLCHGS 283
 113 MNSITSTVDRIIVQGIASELVIPSERHYRNGQWRSVETSGVAEEATSLVWMLCHGS 172

RESULT 8

US-09-769-066-26
 : Sequence 26, Application US/09769065
 : Patent No. US20020107360A1
 : GENERAL INFORMATION:
 : APPLICANT: Fuerst, Thomas R.
 : McAtee, C. Patrick
 : Yarbough, Patricia O.
 : Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-09-769-066
 FILING DATE: 24-Jan-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/542,634
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/CKET NUMBER: 4000-0293.30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:
 LENGTH: 540 amino acids
 TYPE: amino acid
 STRANDEDNESS: Hepatitis E virus (Mexico strain)
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 US-09-769-066-26

Query Match 18.2% Score 120; DB 10; Length 540;
 Best Local Similarity 100.0%; Pred. No. 2.2e-101;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPLODQTNTHIMATASNYAYKPVAPATIPYPPVNPVAVGGYALISFWPQTITPTPSVD 223
 Db 53 LPLODQTNTHIMATASNYAYKPVAPATIPYPPVNPVAVGGYALISFWPQTITPTPSVD 112
 QY 224 MNSITSTVDRIIVQGIASELVIPSERHYRNGQWRSVETSGVAEEATSLVWMLCHGS 283
 Db 113 MNSITSTVDRIIVQGIASELVIPSERHYRNGQWRSVETSGVAEEATSLVWMLCHGS 172

RESULT 9

US-09-769-066-16
 : Sequence 16, Application US/09769066
 : Patent No. US20020107360A1
 : GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.
 McAtee, C. Patrick
 Yarbough, Patricia O.
 Zhang, Yifan
 TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-09-769-066
 FILING DATE: 24-Jan-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/542,634
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/CKET NUMBER: 4000-0293.30
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:
 LENGTH: 549 amino acids
 TYPE: amino acid
 STRANDEDNESS: Hepatitis E virus (Mexico strain)
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 SEQUENCE DESCRIPTION: SEQ ID NO: 16:
 US-09-769-066-16

Query Match 18.2% Score 120; DB 10; Length 549;
 Best Local Similarity 100.0%; Pred. No. 2.2e-101;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPLODQTNTHIMATASNYAYKPVAPATIPYPPVNPVAVGGYALISFWPQTITPTPSVD 223
 Db 53 LPLODQTNTHIMATASNYAYKPVAPATIPYPPVNPVAVGGYALISFWPQTITPTPSVD 112
 QY 224 MNSITSTVDRIIVQGIASELVIPSERHYRNGQWRSVETSGVAEEATSLVWMLCHGS 283
 Db 113 MNSITSTVDRIIVQGIASELVIPSERHYRNGQWRSVETSGVAEEATSLVWMLCHGS 172

RESULT 10

US-09-769-066-14
 : Sequence 14, Application US/09769066
 : Patent No. US20020107360A1
 : GENERAL INFORMATION:
 : APPLICANT: Fuerst, Thomas R.
 : McAtee, C. Patrick
 : Yarbough, Patricia O.
 : Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Ave., Suite 250
 CITY: Palo Alto

STAFF: CA
COUNTRY: USA
ZIP: 94066
COMPUTER: IBM pc compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: pc dos/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: UNKNOWN
PROB APPLICATION DATA:
APPLICATION NUMBER: 09/769,066
FILING DATE: UNKNOWN
ATTORNEY/AGENT INFORMATION:
NAME: Fadian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/AGENT NUMBER: 4,000,029,430
TELEPHONE: (415) 424-0860
TELEFAX: (415) 424-0960
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
STRANDNESS: Right (5' to 3')
TOPOLOGY: Linear
M LEU: 1
HYDROPHILIC: NO
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US 09/769,066, 14

Query Match 18.2% Score 120; db 10; Length 660;
Best Local Similarity 100.0%; Prod. No. 2.7e-101;
Matches 61; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 LPLGLGCHIMALEASNYAYRVAVATDYRGLVNAVGVAGVSPWQTITPSVD 223
DB 144 LPLGLGCHIMALEASNYAYRVAVATDYRGLVNAVGVAGVSPWQTITPSVD 223
QY 224 MNSITSTAVRLVAGVGLASVIVSPHYRNQWSPVTSVAEEFATSLGLMCTHGS 283
DB 224 MNSITSTAVRLVAGVGLASVIVSPHYRNQWSPVTSVAEEFATSLGLMCTHGS 283

RESULT 13
US 09-468-147-175
Sequence 175, Application US/09/468,147A
Publication No. US20040049601A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Erker, James G.
APPLICANT: Dawson, George J.
APPLICANT: Mushahwar, I. K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
FILE REFERENCE: 6242 US 01
CURRENT APPLICATION NUMBER: US/09/468,147A
FILING DATE: 1999-12-21
EARLIER APPLICATION NUMBER: US 09/173,141
EARLIER FILING DATE: 1998-10-15
EARLIER FILING DATE: 1997-10-15
NUMBER OF SEQ ID NOS: 258
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 14
LENGTH: 660
TYPE: PRT

Query Match 9.2% Score 61; db 11; Length 427;
Best Local Similarity 100.0%; Prod. No. 1.7e-471;
Matches 61; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 465 GRCALTEFNALHLEHEDTEETSSAWEHYYEFVAVCHTIVKIVTSVNAWDEK 424
DB 465 GRCALTEFNALHLEHEDTEETSSAWEHYYEFVAVCHTIVKIVTSVNAWDEK 424
QY 425 1 425
DB 92 1 92
RESULT 14
US 09-468-147-176
Sequence 176, Application US/09/468,147A
Publication No. US20040049601A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Schindler, George G.

ORGANISM: Hepatitis E Virus
FEATURES:
OTHER INFORMATION: Xaa Unknown or other at position 494
OTHER INFORMATION: Xaa Unknown or other at position 542
OTHER INFORMATION: Xaa Unknown or other at position 595
US 09-468-147-175

Query Match 11.4% Score 29; db 11; Length 400;

Best Local Similarity 100.0%; Prod. No. 4.7e-60;
Matches 76; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 KATTEVTEVPRAVNAVATETWEETETETVVERSTETTHAVTEWPGTASHVIV 438
DB 199 KATTEVTEVPRAVNAVATETWEETETETVVERSTETTHAVTEWPGTASHVIV 438

QY 249 ERLHYRNQWWSVEI 258
DB 249 ERLHYRNQWWSVEI 258

RESULT 12

US 09-468-147-175

Sequence 175, Application US/09/468,147A

Publication No. US20040049601A1

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories

APPLICANT: Erker, James G.

APPLICANT: Dawson, George J.

APPLICANT: Mushahwar, I. K.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING

FILE REFERENCE: 6242 US 01

CURRENT APPLICATION NUMBER: US/09/468,147A

EARLIER FILING DATE: 1999-12-21

EARLIER APPLICATION NUMBER: 93/097193, 141

EARLIER FILING DATE: 1998-10-15

EARLIER APPLICATION NUMBER: US 09/061,199

NUMBER OF SEQ ID NOS: 258

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO: 175

LENGTH: 427

TYPE: PRT

ORGANISM: Hepatitis E Virus

FEATURES:

OTHER INFORMATION: US 1 S33

OTHER INFORMATION: Xaa Unknown or other at position 148

OTHER INFORMATION: Xaa Unknown or other at position 209

OTHER INFORMATION: Xaa Unknown or other at position 252

US 09-468-147-175

Query Match 9.2% Score 61; db 11; Length 427;

Best Local Similarity 100.0%; Prod. No. 1.7e-471;
Matches 61; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 465 GRCALTEFNALHLEHEDTEETSSAWEHYYEFVAVCHTIVKIVTSVNAWDEK 424
DB 465 GRCALTEFNALHLEHEDTEETSSAWEHYYEFVAVCHTIVKIVTSVNAWDEK 424

QY 425 1 425
DB 92 1 92

RESULT 14

US 09-468-147-176

Sequence 176, Application US/09/468,147A

Publication No. US20040049601A1

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories

APPLICANT: Schindler, George G.

```

1 : APPLICANT: Erker, James G.
2 : APPLICANT: Desai, Suresh M.
3 : APPLICANT: Dawson, George J.
4 : APPLICANT: Mushahwar, I. K.
5 : TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
6 : TITLE OF INVENTION: HEPATITIS E VIRUS
7 : FILE REFERENCE: 6232.US.P1
8 : CURRENT APPLICATION NUMBER: US-09/468,147A
9 : CURRENT FILING DATE: 1999-12-21
10 : EARLIER APPLICATION NUMBER: US-09/173,141
11 : EARLIER FILING DATE: 1998-10-15
12 : EARLIER APPLICATION NUMBER: US-60/061,199
13 : EARLIER FILING DATE: 1997-10-15
14 : NUMBER OF SEQ ID NOS: 258
15 : SOFTWARE: FastSeq for Windows Version 3.0
16 : SEQ ID NO 176
17 : LENGTH: 427
18 : TYPE: PRT
19 : ORGANISM: Hepatitis E Virus
20 : FEATURE:
21 : OTHER INFORMATION: US-2 SG3
22 : OTHER INFORMATION: Xaa = unknown or other if position 114
23 : US -09-468-147-176

Query Match:          9.2%   Score 61; DB 11; Length 327;
Best Local Similarity    100%; Pred.No 1,7e+47;
Matches      61; Conservative     0; Mismatches     0; Indels     0; Gaps     0;

QY       365 GGCTATTTCFNLTLLGCLPTELISSAGGLFYSPVVSANGPTVKIYTSVFNAQQFFG 424
DB        |||||
           42 GGCTATTTCFNLTLLGCLPTELISSAGGLFYSPVVSANGPTVKIYTSVFNAQQFFG 91

QY             425 1 425
DB              92 | 92

RESULT 14
US-09-468-147-199
Sequence 199, Application US-09/468147A
Publication No. US20030049601A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Schlauder, George C.
APPLICANT: Erker, James G.
APPLICANT: Desai, Suresh M.
APPLICANT: Dawson, George J.
APPLICANT: Mushahwar, I. K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
TITLE OF INVENTION: HEPATITIS E VIRUS
FILE REFERENCE: 6232-US.P1
CURRENT APPLICATION NUMBER: US-09/468,147A
CURRENT FILING DATE: 1999-12-21
EARLIER APPLICATION NUMBER: US-09/173,141
EARLIER FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: US-60/061,199
EARLIER FILING DATE: 1997-10-15
NUMBER OF SEQ ID NOS: 258
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 199
LENGTH: 438
TYPE: PRT
ORGANISM: Hepatitis E Virus
FEATURE:
OTHER INFORMATION: cksof2m-2.pep
US-09-468-147-199

Query Match            9.2%   Score 61; DB 11; Length 338;
Best Local Similarity   100%; Pred.No 1,8e+47;
Matches      61; Conservative     0; Mismatches     0; Indels     0; Gaps     0;

QY       465 GGCTATTTCFNLTLLGCLPTELISSAGGLFYSPVVSANGPTVKIYTSVFNAQQFFG 424

```


GenInfo version 5.1.4

Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 09:41:04 : Search time 24 seconds
(without alignments)
2595.423 Million cell updates/s/cp

Title: US-09-851-410a-8
Protein source: 660
Sequence: 1 MRPRTVILLILLMLPLPAP.....GSTVAELQPLKMKVGKTRRL 660

Scoring table: cd100

Gapop 60.0 : Gapext 60.0

Searches: 283508 seqs, 9616082 residues

Word size: 3

Total number of hits satisfying chosen parameters: 283508

Minimum DB Seq length: 0

Maximum DB Seq length: 2000000000

Post processing: Last 100 First 45 summaries

Database: 1

1: p111*

2: p112*

3: p113*

4: p114*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	660	1 VHW02	structural protein
2	120	18.2	659	1 B41212	structural protein
3	9	1.4	1091	2 F85928	hypothetical prote
4	8	1.2	41	2 B8208	proteinase 11 - bla
5	6	1.2	172	2 F87649	ExoD/Tox family p
6	6	1.2	199	2 G55539	photosystem I, ox
7	8	1.2	252	2 F72224	oxygen evolving en
8	8	1.2	246	2 G97489	hypothetical prote
9	8	1.2	243	1 W088K	U13 protein - huma
10	8	1.2	240	2 F72797	hypothetical prote
11	8	1.2	241	2 S7170	ribosomal protein
12	8	1.2	252	2 F15737	hypothetical prote
13	8	1.2	273	2 F48639	hypothetical prote
14	8	1.2	272	2 A8794	thiosulfate sulfur
15	8	1.2	292	2 H70313	cochlearin synthet
16	8	1.2	508	2 S68376	aryl hydrocarbon r
17	8	1.2	412	2 F86563	protein 10355.1 (1
18	8	1.2	447	2 F91678	GTP binding protei
19	8	1.2	447	2 G71521	probable GTP bindi
20	8	1.2	448	2 G83938	hypothetical prote
21	8	1.2	507	2 G01614	zinc finger protei
22	8	1.2	666	2 F22943	hypothetical prote
23	8	1.2	678	2 A75580	hypothetical prote
24	8	1.2	776	2 A5438	AB receptor nucle
25	8	1.2	789	2 F95550	aryl hydrocarbon r
26	8	1.2	793	2 A56231	aryl hydrocarbon r
27	8	1.2	805	2 F87635	aryl hydrocarbon r
28	8	1.2	793	2 A40437	probable haemoco
29	8	1.2	233	2 F25410	hypothetical prote

ALIGNMENTS

RESULT 1

VHW02

structural protein 2 precursor

Species: hepatitis E virus

Date: 30 Sep 1992 #sequence revision: 39 Sep 1992 #1st change: 23 Jul 1999

Accession: C40778

Rfam, A.W.: Smith, M.M.; Ghetti, M.; Bianchi, J.C.; Rodley, L.W.; Fry, K.E.; Rogers, J.

Virology 185, 120-131, 1991

A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full length

A:Reference number: A40778; M0110; G024967; F0110; F02970

Accession: C40778

Molecule type: genomic RNA

Residues: 1-660 - IAM

Accession: G007428; M0110; G024967; F0110; F02970

A:Note: The authors translated the open ORF for residue 2 as Ala

Support family: hepatitis E virus structural protein 2

Keywords: structural protein

F:1-222nucleic acid sequence; #status predicted: S03

F:23-660protein; structural protein 2; #status predicted: S02

Query Match

100.0% Score 660; ID 1; Length 660

Best local similarity 100.0% Prod. No. 0

Matches 660; Conservative 0; Mismatches 0; Gaps 0;

Q9

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

C: Genet test

A: Gene: pda4
 A: Species: chloroplast; membrane protein; photosynthesis; photosystem II
 A: Note: host Homo sapiens (hum)
 C: Date: 11 Oct 1999 #sequence_revision 40 Sep 2001 #text_change 18 Nov 2002
 C: Accession: S72224

Query Match 1.2% Score 8; DB 2; Length 199;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 RFAVASR 100
 ID 1111111
 ID 11 RFAVASR 17

RESULT 7

S72224
 A: Species: Volvox carteri
 A: Note: 04 Feb 1998 #sequence_revision 13 Feb 1998 #text_change 21 Jul 2000
 C: Accession: S72224
 A: Title: Three abundant germ line specific transcripts in Volvox carteri encode photosyn
 A: Reference number: S72222; MIMD:5676459; EMBL:878179
 A: Accession: S72224
 A: Status: preliminary; nucleic acid sequence not shown
 A: Molecule type: mRNA
 A: Residues: 1-202 (M)
 A: Cross references: EMBL:022439; NID:0732545; FID:AA049800.1; FID:Q732546
 C: Genet test
 A: Genes: nucleat
 A: Species: chloroplast

Query Match 1.2% Score 8; DB 2; Length 202;
 Best local similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 RFAVASR 100
 ID 1111111
 ID 10 RFAVASR 17

RESULT 8

S72486
 A: Species: Arabidopsis thaliana
 A: Note: 09 Sep 2001 #sequence_revision 40 Sep 2001 #text_change 18 Nov 2002
 C: Accession: S72486
 A: Title: B1.1 Hinkley, G.; Gattang, S.; Miller, N.; Blanchard, M.; Gourollo, B.; Goldman
 A: Reference number: S72486; MIMD:5676459; EMBL:878179
 A: Accession: S72486
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-226 (M)
 A: Cross references: GBAF007869; P10N:AAK86860.1; P10N:Q15156076; GSPDB:GN50169
 C: Genet test
 A: Genes: Asp 1942
 A: Map position: circular chromosome

Query Match 1.2% Score 8; DB 2; Length 226;
 Best local similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 RFAVASR 100
 ID 1111111
 ID 20 RFAVASR 17

RESULT 9

S72486

WIDEHR

OLA protein = human herpesvirus 2 (strain Hb22)
 A: Species: human herpesvirus 2
 A: Note: host Homo sapiens (hum)
 C: Date: 11 Oct 1999 #sequence_revision 40 Sep 2001 #text_change 18 Nov 2002
 C: Accession: J01496
 A: Title: Comparison of sequence analysis of the long repeat regions and adjoining parts
 A: Reference number: J01494; MIMD:5676459; EMBL:166247
 A: Accession: J01496
 A: Molecule type: DNA
 A: Residues: 1-243 (M)
 A: Cross references: GBAF007869; MIMD:5676459; FID:AA049800.1; FID:Q732546
 C: Genet test
 A: Genes: HLA
 C: Superfamily: viral; herpes virus; gene 6; protein

Query Match 1.2% Score 8; DB 2; Length 243;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 141 SMLSSVA 148
 ID 1111111
 ID 75 SMLSSVA 82

RESULT 10

T24797
 A: Species: human herpesvirus 2 (strain Hb22)
 A: Note: 15 Oct 1999 #sequence_revision 40 Sep 2001 #text_change 18 Nov 2002
 C: Accession: T24797
 A: Title: Comparison of sequence analysis of the long repeat regions and adjoining parts
 A: Reference number: T24794; MIMD:5676459; EMBL:166247
 A: Accession: T24797
 A: Molecule type: DNA
 A: Residues: 1-240 (M)
 A: Cross references: GBAF007869; MIMD:5676459; FID:AA049800.1; FID:Q732546
 C: Genet test
 A: Genes: MIMD:5676459
 A: Map position: 4
 A: Status: preliminary; translated from GBAF007869

Query Match 1.2% Score 8; DB 2; Length 240;
 Best local similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 42 RSGGSRG 49
 ID 45 RSGGSRG 52

RESULT 11

S74170
 A: Species: chloroplast; porphyra purpurea
 A: Note: 19 Mar 1997 #sequence_revision 40 May 1997 #text_change 18 Nov 2002
 C: Accession: S74170
 A: Title: Complete nucleotide sequence of the porphyra purpurea chloroplast genome
 A: Reference number: S74169
 A: Accession: S74170
 A: Status: nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A: Residues: 1-241 (M)
 A: Cross references: EMBL:008604; NID:0732545; FID:AA049800.1; FID:Q732546
 C: Genet test
 A: Note: the nucleotide sequence was submitted to the EMBL Data Library October 1995


```

QY 181 NYAYRYVARATIPYVETLV:NAVZGVAISLSHWQTTTTTTSVDMNSITSTVEILLVSPGI 240
DB 181 NYAYRYVARATIPYVETLV:NAVZGVAISLSHWQTTTTTTSVDMNSITSTVEILLVSPGI 240
QY 241 ASELVTPSEPLHPNCGWPSVETSGVAPPEATSGVMVLCIHGSLVNSVNTPTPYTGALGL 300
DB 241 ASELVTPSEPLHPNCGWPSVETSGVAPPEATSGVMVLCIHGSLVNSVNTPTPYTGALGL 300
QY 301 DFALELEFRLNCTPGNTNTVPSVSYSTARHPLRPGAGTAEIITTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRLNCTPGNTNTVPSVSYSTARHPLRPGAGTAEIITTAATRFMKDLYFTSTNG 360
QY 361 VSEFGECIALTPNADILGGLPTPELLISSANQQLFSPVVSANGCEPTVKLYTSVENAQ 420
DB 361 VSEFGECIALTPNADILGGLPTPELLISSANQQLFSPVVSANGCEPTVKLYTSVENAQ 420
QY 421 QDRGIAIPHDIDIGESKVVYQYDNGHQDRPTSPAPSRPFSVLKANDVWLWLSLTAAY 480
DB 421 QDRGIAIPHDIDIGESKVVYQYDNGHQDRPTSPAPSRPFSVLKANDVWLWLSLTAAY 480
QY 481 DQSTYSSSGPVVYSVSVTLVNVATGACAVAPSLDWTKVTLGDRPLSTIQOYSKTFEVL 540
DB 481 DQSTYSSSGPVVYSVSVTLVNVATGACAVAPSLDWTKVTLGDRPLSTIQOYSKTFEVL 540
QY 541 LRKLSFWPAGTICAGYPYNYNTASDQLLVENACHKVAISLYTISGACGWSISAVAV 600
DB 541 LRKLSFWPAGTICAGYPYNYNTASDQLLVENACHKVAISLYTISGACGWSISAVAV 600
QY 601 LAPHSALALERTLVPAPAHPTDDPCPRPLSGAGAFQSTVAELQFLMKVKGTREL 660
DB 601 LAPHSALALERTLVPAPAHPTDDPCPRPLSGAGAFQSTVAELQFLMKVKGTREL 660

RESULT 2
VST2_HEVPA STANDARD: PRI: 660 AA
AC P3426;
DT 01-FEB-1994 (Rel. 28, Created)
DI 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2115700; PubMed=1741327;
RA Tsarev S.A., Enersen S.G., Peyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
PT "Characterization of a prototype strain of hepatitis E virus."
PI Proc Natl Acad Sci U S A 89:553-553(1992)
CC - FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/industry
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M80581; AAA45727.1;
CC InterPro: IPR004261; SP2.
CC Pfam: PF03014; SP2; 1.
CC
CC Signal.
CC CHAIN 1 22 BY SIMILARITY.
CC FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
CC SEQUENCE 660 AA. 70580 MW. 80555.5 kDa.

```

```

Query Match 61.8% Score 408: DB 1: Length 660:
Basic Local Similarity 69.9% Prof. No. 0:
Matches 508: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
QY 97 ASRRRTTAAAPLJAVAFADHTPPVVDSSRGALLRQYNLSISTSSVAIGINIVLY 154
DB 97 ASRRRTTAAAPLJAVAFADHTPPVVDSSRGALLRQYNLSISTSSVAIGINIVLY 154
QY 157 AAPLSPLLPLOQGTNTHIMATEASNYAYRYVARATIPYVETLV:NAVZGVAISLSHWQ 214
DB 157 AAPLSPLLPLOQGTNTHIMATEASNYAYRYVARATIPYVETLV:NAVZGVAISLSHWQ 214
QY 217 TPTTSVDMNSITSTVEILLVSPGI:NAVZGVAISLSHWQTTTTTTSVDMNSITSTVE 276
DB 217 TPTTSVDMNSITSTVEILLVSPGI:NAVZGVAISLSHWQTTTTTTSVDMNSITSTVE 276
QY 277 MLCIHGSPVNSVNTPTPYTGALGLIHGSLVNSVNTPTPYTGALGLIHGSLVNSVNTPT 346
DB 277 MLCIHGSPVNSVNTPTPYTGALGLIHGSLVNSVNTPTPYTGALGLIHGSLVNSVNTPT 346
QY 337 GTAEIITTAATRFMKDLYFTSTNGVCEIGRGIALTPNADILGGLPTPELLISSANQQL 396
DB 337 GTAEIITTAATRFMKDLYFTSTNGVCEIGRGIALTPNADILGGLPTPELLISSANQQL 396
QY 397 VSEFGECIALTPNADILGGLPTPELLISSANQQLFSPVVSANGCEPTVKLYTSVENAQ 454
DB 397 VSEFGECIALTPNADILGGLPTPELLISSANQQLFSPVVSANGCEPTVKLYTSVENAQ 454
QY 457 ASRRRTTAAAPLJAVAFADHTPPVVDSSRGALLRQYNLSISTSSVAIGINIVLY 514
DB 457 ASRRRTTAAAPLJAVAFADHTPPVVDSSRGALLRQYNLSISTSSVAIGINIVLY 514
QY 517 TPTTSVDMNSITSTVEILLVSPGI:NAVZGVAISLSHWQTTTTTTSVDMNSITSTVE 574
DB 517 TPTTSVDMNSITSTVEILLVSPGI:NAVZGVAISLSHWQTTTTTTSVDMNSITSTVE 574
QY 577 HPVALSTVITSLGAGFVSISAVAVLAPHS 605
DB 577 HPVALSTVITSLGAGFVSISAVAVLAPHS 605
RESULT 4
VST2_HEVNY STANDARD: PRI: 660 AA.
AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DI 01-OCT-1993 (Rel. 27, Last sequence update)
DE 01-NOV-1995 (Rel. 42, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strains Myanmar) (HEV).
OC Viruses; ssRNA positive strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=41769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227574; PubMed 8470471;
RA Aye I.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ishikawa M.,
RA Rikihisa T., Wun K.;
PT "Sequence and gene structure of the hepatitis E virus isolated from
RI Myanmar."
RL Virus Genes 7:55-59(1993)
CC - FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/industry
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M80581; AAA45727.1;
CC InterPro: IPR004261; SP2.
CC Pfam: PF03014; SP2; 1.
CC
CC Signal.
CC CHAIN 1 22 BY SIMILARITY.
CC FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
CC SEQUENCE 660 AA. 70580 MW. 80555.5 kDa.

```

```

DB EMBL: J104002; HA001174.1;
DB InfoProc: J104002.261; SP2:
DB Pfam: PF03014; SP2: 1;
KW SIGNAL; 1 22 BY SIMILARITY;
FT CHAIN 23 660 STRUCTURAL PROTEIN 2;
SQ SEQUENCE 660 AA; 70996 MW; 5AB2AMFA2556253 CRC64;

Query Match
Best local similarity 100.0%; Score 128; DB 1; Length 660;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 1 MRPDPTLLMLLMLPMLPAPRPGQSPRRKSPSGSGPMDGVDSQPAIPYIHTN 60
DB 1 MRPDPTLLMLLMLPMLPAPRPGQSPRRKSPSGSGPMDGVDSQPAIPYIHTN 60
Q7 61 PFADDTAAAGAPRVQRPAPRPGISAWPQAGRPVAVASRRKRTTGAAPLTAVAPADTP 120
DB 1 PFADDTAAAGAPRVQRPAPRPGISAWPQAGRPVAVASRRKRTTGAAPLTAVAPADTP 120
Q7 121 PVPDVDSR 128
DB 121 PVPDVDSR 128
Q7 121 PVPDVDSR 128
DB 121 PVPDVDSR 128

RESULT 4
VST2_HEVME
ID VST2_HEVME STANDARD; PRT; 485 AA;
AC Q00270;
DT 01-OCT-1993 (Ref. 27, Created)
DI 01-OCT-1993 (Ref. 27, Last sequence update)
DI 01-OCT-1993 (Ref. 27, Last annotation update)
DE Structural protein 2 (Fragment)
OS Hepatitis E virus (isolate Rhesus) (HEV)
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses;
OC NCHL TaxID 41769;
FN [1]
SQ SEQUENCE FROM N.A.
MEDLINE 94261477; PubMed 1594074;
RA Bradley D.W., Suzuki K., Harada S., Iida T., Hara T., Kato S.S.,
RA Bradley D.W., Tam A.W., Reyes G.F.;
RT "Molecular cloning and sequencing of the Mexican isolate of hepatitis
E virus (HEV).";
RL Virology 192:556-558(1992)
CC 1 FUNCTION: CONTAINS A HIGH MOLECULAR WEIGHT ACTIN-BINDING SITES
CC THAT IT MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS RNA
CC BY EFFECTIVELY NEUTRALIZING THE REPLICATIVELY INACTIVATED RNA.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.sib.ch).

```

```

DB EMBL: J104002; HA001174.1;
DB InfoProc: J104002.261; SP2:
DB Pfam: PF03014; SP2: 1;
KW SIGNAL; 1 22 BY SIMILARITY;
FT CHAIN 23 660 STRUCTURAL PROTEIN 2;
SQ SEQUENCE 660 AA; 70996 MW; 5AB2AMFA2556253 CRC64;

Query Match
Best local similarity 100.0%; Score 128; DB 1; Length 660;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 1 MRPDPTLLMLLMLPMLPAPRPGQSPRRKSPSGSGPMDGVDSQPAIPYIHTN 60
DB 1 MRPDPTLLMLLMLPMLPAPRPGQSPRRKSPSGSGPMDGVDSQPAIPYIHTN 60
Q7 61 PFADDTAAAGAPRVQRPAPRPGISAWPQAGRPVAVASRRKRTTGAAPLTAVAPADTP 120
DB 1 PFADDTAAAGAPRVQRPAPRPGISAWPQAGRPVAVASRRKRTTGAAPLTAVAPADTP 120
Q7 121 PVPDVDSR 128
DB 121 PVPDVDSR 128
Q7 121 PVPDVDSR 128
DB 121 PVPDVDSR 128

RESULT 4
VST2_HEVME
ID VST2_HEVME STANDARD; PRT; 485 AA;
AC Q00270;
DT 01-OCT-1993 (Ref. 27, Created)
DI 01-OCT-1993 (Ref. 27, Last sequence update)
DI 01-OCT-1993 (Ref. 27, Last annotation update)
DE Structural protein 2 (Fragment)
OS Hepatitis E virus (isolate Rhesus) (HEV)
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses;
OC NCHL TaxID 41769;
FN [1]
SQ SEQUENCE FROM N.A.
MEDLINE 94261477; PubMed 1594074;
RA Bradley D.W., Suzuki K., Harada S., Iida T., Hara T., Kato S.S.,
RA Bradley D.W., Tam A.W., Reyes G.F.;
RT "Molecular cloning and sequencing of the Mexican isolate of hepatitis
E virus (HEV).";
RL Virology 192:556-558(1992)
CC 1 FUNCTION: CONTAINS A HIGH MOLECULAR WEIGHT ACTIN-BINDING SITES
CC THAT IT MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS RNA
CC BY EFFECTIVELY NEUTRALIZING THE REPLICATIVELY INACTIVATED RNA.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.sib.ch).
DB EMBL: M74506; AAA45742.1;
DR PIR: I44212; I44212;
DR InfoProc: J104002.261; SP2:
DR Pfam: PF03014; SP2: 1;
KW SIGNAL;
FT CHAIN 1 22 BY SIMILARITY;
FI CHAIN 23 660 STRUCTURAL PROTEIN 2;
SQ SEQUENCE 660 AA; 70940 MW; 5F6E7AEEF0E9E9A CRC64;

Query Match
Best local similarity 100.0%; Score 120; DB 1; Length 660;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 164 LPLQGGNHHIMAEHAEKAAEAYALATKTHANRAVDAEALATVQTHLITSL 224
DB 164 LPLQGGNHHIMAEHAEKAAEAYALATKTHANRAVDAEALATVQTHLITSL 224
Q7 225 MRSFLATVAVQVQASAAVTSKDYKQKQWVLSVALAASQAMV 284
DB 225 MRSFLATVAVQVQASAAVTSKDYKQKQWVLSVALAASQAMV 284
Q7 224 MRSFLATVAVQVQASAAVTSKDYKQKQWVLSVALAASQAMV 284
DB 224 MRSFLATVAVQVQASAAVTSKDYKQKQWVLSVALAASQAMV 284

RESULT 6
PSBQ_CHLRE
ID PSBQ_CHLRE STANDARD; PRT; 199 AA;
AC P12852;
DI 01-OCT-1989 (Ref. 12, Created)
DI 01-OCT-1989 (Ref. 12, Last sequence update)
DI 16-OCT-2001 (Ref. 40, Last annotation update)
DE oxygen-evolving enhancer protein 3, chloroplast precursor (oPEP3)
GN PSBQ;
OS Chlamydomonas reinhardtii;
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadales; Chlamydomonadales;
OC NCHL TaxID 4055;
FN [1]
SQ SEQUENCE FROM N.A.
MEDLINE 14767071; PubMed 14767071;
RA Mayfield S.H.;
RT "Analysis of the genes of the oPEP1 and oPEP2 proteins of the

```

RI photosystem II complex of Chlamydomonas reinhardtii.
 RL Plant Mol. Biol. 12:492-493(1994).
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch>)
 or send an email to license@isb-sib.ch.
 CC
 DR EMBL: X13832, CAA32091.1.
 DR PIR: S05509; S05509.
 KW photosynthesis; Photosystem II; Chloroplast; Transit peptide;
 FW Thylakoid; Membrane.
 FT TRANSIT 51 CHLOROPLAST.
 FT CHAIN 52 190 OXYGEN-EVOLVING ENHANCER PROTEIN 4.
 SQ SEQUENCE 190 AA: 21824 MW: 30767733407066 cpo64.

 Query Match 1.28; Score 8; DB 1; Length 199;
 Best Local Similarity 100.00; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 93 PPAVASPP 100
 DB 10 RPAVASRR 17
 |||||
 RESULT 7
 ID PSBO-VOLCA STANDARD; PPT; 202 AA.
 AC Q41643;
 DT 16-OCT-2001 (Rel. 40, Created)
 DI 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oxygen-evolving enhancer protein in 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000.
 CN PSBO.
 OS Volvox carteri.
 OR Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales.
 OC Volvocaceae; Volvox.
 OX NCBI_TaxID=3067;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=f, Naqariensis / HK10;
 RX MEDLINE=95376499; PubMed=8781179;
 RA Choi G., Przybylska M., Straus D.;
 RI "Three abundant germ line-specific transcripts in Volvox carteri
 encode photosynthetic proteins."
 RL Curr. Genet. 40:347-355(1996)
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch>)
 or send an email to license@isb-sib.ch.
 CC
 DR EMBL: U22330; AAR40980.1;
 DR PIR: S72224; S72224.
 DR TIGRGRAMS: TIGR01409; TAT_signal_seq: 1.
 KW photosynthesis; Photosystem II; Chloroplast; Transit peptide;
 FW Thylakoid; Membrane.
 FT TRANSIT 1 54 CHLOROPLAST (POTENTIAL).
 FT CHAIN 55 202 OXYGEN-EVOLVING ENHANCER PROTEIN 7.
 SQ SEQUENCE 202 AA: 22290 MW: 3103253784023180 cpe64.

Query Match 1.28; Score 8; DB 1; Length 202;
 Best Local Similarity 100.00; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 93 RPAVASRR 100
 DB 10 RPAVASRR 17
 |||||
 RESULT 8
 ID PSBO-VOLCA STANDARD; PPT; 202 AA.
 AC Q41643;
 DT 16-OCT-2001 (Rel. 40, Created)
 DI 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oxygen-evolving enhancer protein in 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000.
 CN PSBO.
 OS Volvox carteri.
 OR Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales.
 OC Volvocaceae; Volvox.
 OX NCBI_TaxID=3067;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=f, Naqariensis / HK10;
 RX MEDLINE=95376499; PubMed=8781179;
 RA Choi G., Przybylska M., Straus D.;
 RI "Three abundant germ line-specific transcripts in Volvox carteri
 encode photosynthetic proteins."
 RL Curr. Genet. 40:347-355(1996)
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch>)
 or send an email to license@isb-sib.ch.
 CC
 DR EMBL: U22330; AAR40980.1;
 DR PIR: S72224; S72224.
 DR TIGRGRAMS: TIGR01409; TAT_signal_seq: 1.
 KW photosynthesis; Photosystem II; Chloroplast; Transit peptide;
 FW Thylakoid; Membrane.
 FT TRANSIT 1 54 CHLOROPLAST (POTENTIAL).
 FT CHAIN 55 202 OXYGEN-EVOLVING ENHANCER PROTEIN 7.
 SQ SEQUENCE 202 AA: 22290 MW: 3103253784023180 cpe64.


```

RP SEQUENCE FROM N.A.
RA Smith D.P., Robinson K.,
RI Submitted (SEP-1994) to the EMRL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TN:
RX MEDLINE=2119732; PubMed=11234002;
RA Cole S.T., Elmqvist K., Parkhill J., James K.D., Thomson N.P.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor K.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hurrest T., Jais K., Jais K., Jais K., Jais K., Jais K.,
RA Murphy J., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Butler S., Seeger K., Simon S., Simmonds M., Skellern J., Squares F.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Harrell B.G.;
RI "Massive gene decay in the leprosy bacillus."
RL Nature 406:1007-1011(2003)
CC -!- FUNCTION: MAY BE A SULFOTRANSFERASE INVOLVED IN THE FORMATION OF
CC THIOSULFATE
CC -!- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.
CC -!- SIMILARITY: Contains 2 rhodanese domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMRL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC
CC EMBL: U15182; AAA62982.1;
DR EMBL: AL584924; CAC31153.1;
DR PIR: A87184; A87184.
DR Leprosoma; ML2198;
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR001307; Rhodanese
DR Pfam: PF00581; Rhodanese; 2.
DR SMART: SM00450; PHOD; 2.
DR PROSITE: PS00683; PHODANESF_2; 1.
DR PROSITE: PS00206; PHODANESF_3; 2.
KW Transferrase; Complete proteome; Repeat.
FT DOMAIN 18 125 RHODANESF 1.
FT DOMAIN 154 274 RHODANESF 2.
FT ACT_SITE 233 233 BY SIMILARITY.
FT ACT_SITE 277 277 AA: 527E05C9FF93969D CPC64;
SQ SEQUENCE 277 AA: 527E05C9FF93969D CPC64;
Query Match 1.28; Score 8; DB 1; Length 277;
Best Local Similarity 100.0%; Pred No. R 7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 521 IICPPLST 528
DB 118 LDCRPLST 125
|||||||
|||||||
RESULT 12
YNP1_CAEEL STANDARD; PRT; 412 AA.
AC P34554;
DT 01-FEB-1994 (Rel. 28, Created)
DI 30-MAY-2000 (Rel. 39, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein T05G5.1 in chromosome III precursor.
GN T05G5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Chordata; Phlebotomidae; Phlebotomidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;

```

```

KA Wilson R., Alston R., Anderson K., Barnes C., Berks M.,
KA Benfield J., Burton J., Connel M., Copsey L., Cooper J., Gouison A.,
KA Craxton M., Dear S., De Z., Durbin P., Pavolito A., Fraser A.,
KA Frailton L., Gardner A., Green P., Hawkins T., Hillier L., Jior M.,
KA Johnston L., Jones M., Kershaw J., Kirsten J., Laisner N.,
KA Lattelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
KA Parsons J., Pary C., Pikel L., Pikel A., Sandberg P., Shownkeen R.,
KA Sims M., Smallwood N., Smith A., Smith M., Smitlammer F., Staden R.,
KA Sultun J., Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K.,
KA Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
KA Wolfenden P.;
PT "2.3 Mb of sequencing overlap the sequence from chromosome III of C.
EL elegans."
RL Nature 423:32-38(1994).
RN [2]
RP REVISIONS.
RA Durbin R.;
RA Durbin R.;
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMRL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC
CC EMBL: Z27079; CAAL594.1;
DR PIR: FR8563; FR8563.
DR HSSP: P01132; LEFG.
DR WormPep: T05G5.1; CE21151.
DR InterPro: IPR006210; IEGF.
DR SMART: SM00181; EGF; 1.
KW Hypothetical protein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 412 HYPOTHETICAL 49.5 KDa PROTEIN T05G5.1.
SQ SEQUENCE 412 AA: 49520 MW: 41A4D68C151918BA CR764;
Query Match 1.44; Score 8; DB 1; Length 412;
Best Local Similarity 100.0%; Pred No. R 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ILILLMF 13
|||||||
DB 5 ILILLMF 12
|||||||
RESULT 14
MER3_HUMAN STANDARD; PRT; 507 AA.
ID MER3_HUMAN
AC Q13064;
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Makrin 3 (Zinc finger protein 127).
GN MKRN3 OR ZNF127.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE 93211018; PubMed 10196367;
RA Jung M.F.-C., Gray T.A., Ji Y., Glenn C.G., Saitoh S., Driscoll D.J.,
RA Nicholls R.D.;
RT "A novel imprinted gene, encoding a RING zinc finger protein, and
RT overlapping antisense transcript in the Prader-Willi syndrome
RT critical region."
RL Hum. Mol. Genet. 8:783-793(1999).
CC -!- TISSUE SPECIFICITY: ubiquitous.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 3 CCH1-type zinc fingers.
CC

```


Matches	197;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	436	DATAELTTAATREMKDIYFTSTNGVGRIGRGTAITLFLNLAPTLLGSLPTELISSAGGQL	495						
Db									
QY	446	DATAELTTAATREMKDLYFTSINGVGEIGRGIALTLFLNLADTLGLGLPTELISSAGGQL	395						
Db									
QY	396	FYSRPVVSANGPEPIVKLYTSVENAQKKGIAIPHDIDLGSRVVIQDYDNQHEQDRPTPS	455						
Db									
QY	396	FYSRPVVSANGPEPIVKLYTSVENAQKKGIAIPHDIDLGSRVVIQDYDNQHEQDRPTPS	455						
Db									
QY	456	PAPSRPFSLRANVLWLSLTAARYDQSTYGSSTGPVYVSDSVILNVNVAITGACAVARSLD	515						
Db									
QY	456	PAPSRPFSLRANVLWLSLTAARYDQSTYGSSTGPVYVSDSVILNVNVAITGACAVARSLD	515						
Db									
QY	516	WTKVTLDGRPLSTIQY	532						
Db									
QY	516	WTKVTLDGRPLSTIQY	532						
Db									

Search completed: August 5, 2003, 09:33:41
Job time : 48 secs

60953 GR. Yarbough pub. Beadley LW. Krawczynski KZ. Tam AW. Fry KE;
 XX WFL: 2001: 442705/46.
 DR N PSUB: AAF94495.
 XX
 XX New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)
 ET hepatitis viral agent, useful in diagnosing infection by an enterically
 transmitted agent (e.g. ET NANB virus), as well as in vaccine
 production
 XX
 XX Disclosure: Columns 61-68: 45pp; English.
 XX
 XX the invention relates to an isolated DNA comprising the genome of an
 enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also
 referred as HEV). The DNA sequences of their fragments are useful in
 preparing ET NANB viral proteins and as probes for virus detection.
 XX These are particularly useful in diagnosing infection by an enterically
 transmitted agent (e.g. ET NANB virus), as well as in vaccine production.
 XX the present sequence represents a ORF2 protein encoded by a ET-NANB viral
 DNA sequence from HEV-Burma strain
 XX
 XX 35-epi-acc 660 AA;

Query Match 100.0% Score 4406; DB 22; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0.66 263;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

07 1 MKRPFLLLLMLFAPPGQSGRRGRSGSGGFWGDRVDSQPPALFYDPTN 60
 10 1 MKRPFLLLLMLFAPPGQSGRRGRSGSGGFWGDRVDSQPPALFYDPTN 60
 07 61 PRAPVTAAGAGPVRGAFARLGSAMWDQAGRTAVASRRPPTTAGAAPTAVAPADTP 120
 10 61 PRAPVTAAGAGPVRGAFARLGSAMWDQAGRTAVASRRPPTTAGAAPTAVAPADTP 120
 07 121 PWGDSRGLLRGYNSISGLSSVATGNTNVIYAAPLSPILPLQGTNHHMALEAS 180
 10 121 PWGDSRGLLRGYNSISGLSSVATGNTNVIYAAPLSPILPLQGTNHHMALEAS 180
 07 181 NYAGYVAPATIRYPLVNAVGGYALISFWPQTTTPTTSVMNSSTSTAVRLVQPGI 240
 10 181 NYAGYVAPATIRYPLVNAVGGYALISFWPQTTTPTTSVMNSSTSTAVRLVQPGI 240
 07 241 ASLLVTPSRHYRNGWRKSVLSVAFPEATSLVLC HGSIVNSYNTPTPGALGLL 300
 10 241 ASLLVTPSRHYRNGWRKSVLSVAFPEATSLVLC HGSIVNSYNTPTPGALGLL 300
 07 301 DEALELEFKNLTPGNTIRVSKYSSIAHRLRPGAGTAEITTTAATPFMKOLYFTSTNC 360
 10 301 DEALELEFKNLTPGNTIRVSKYSSIAHRLRPGAGTAEITTTAATPFMKOLYFTSTNC 360
 07 361 VRIIGGIALTFNLADTLGLLPTLSSAGGGLFYSGRWVVSANRPPTVKI YTSVENAQ 420
 10 361 VRIIGGIALTFNLADTLGLLPTLSSAGGGLFYSGRWVVSANRPPTVKI YTSVENAQ 420
 07 421 LKQKLSFWACATTKACYPYNNNTASTGLLVENAGHEVALSTYTTSLGAGPWSISAVV 480
 10 421 LKQKLSFWACATTKACYPYNNNTASTGLLVENAGHEVALSTYTTSLGAGPWSISAVV 480
 07 481 LKQKLSFWACATTKACYPYNNNTASTGLLVENAGHEVALSTYTTSLGAGPWSISAVV 540
 10 481 LKQKLSFWACATTKACYPYNNNTASTGLLVENAGHEVALSTYTTSLGAGPWSISAVV 540
 07 541 LKQKLSFWACATTKACYPYNNNTASTGLLVENAGHEVALSTYTTSLGAGPWSISAVV 600
 10 541 LKQKLSFWACATTKACYPYNNNTASTGLLVENAGHEVALSTYTTSLGAGPWSISAVV 600
 07 601 LAHSAALALLELLYPAFAHIFDQPTPTCPPLGLQYAFQSTVAELQPLMKVSKTPEL 660
 10 601 LAHSAALALLELLYPAFAHIFDQPTPTCPPLGLQYAFQSTVAELQPLMKVSKTPEL 660

RES01.1 9

AAL015699
 ID AAL015699 standard; Protein: 660 AA
 XX
 XX AAL015699;
 XX
 XX 06 NOV 2002 (First entry)
 XX
 XX Hepatitis E virus (Burma strain) ORF2 protein.
 XX
 XX HEV; enterically transmitted nonA/nonB hepatitis viral agent;
 XX Burma strain; EHEC ORF2.
 XX
 XX Hepatitis E virus.
 XX
 XX US6479891-B1.
 XX
 XX 40 APR 2002.
 XX
 XX 19 APR 2000; 2000US 055424.
 XX
 XX 25 JUN 1994; 94US 0279824.
 XX
 XX 07 JUN 1995; 95US 0478507.
 XX
 XX 15 APR 1991; 91US 0681078.
 XX
 XX 17 JUN 1988; 88US 0208977.
 XX
 XX 11 APR 1989; 89US 0346672.
 XX
 XX 16 JUN 1989; 89US 0367486.
 XX
 XX 13 OCT 1989; 89US 0420921.
 XX
 XX 05 APR 1990; 90US 0505888.
 XX
 XX (ORF2) OF THE HEPATITIS E VIRUS GENOMES.
 XX (ORF2) GENOMES OF THE HEPATITIS E VIRUS.
 XX
 XX Reyes GR. Yarbough pub. Beadley LW. Krawczynski KZ. Tam AW. Fry KE;
 XX
 XX WFL: 2002: 517277/55.
 XX N PSUB: AAL56886.
 XX
 XX Detecting the presence of enterically transmitted nonA/nonB hepatitis
 PT viral (HEV) agent in bile samples from infected humans and monkeys
 PT using polymerase chain reaction
 XX
 XX Disclosure: Column 19-24: 61pp; English.
 XX
 XX The invention comprises a method for detecting the presence of
 XX enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a
 XX sample and isolated HEV agents or nucleic acid fragments produced by the
 XX agent. The method utilizes PCR using bile from a human or cynomolgus
 XX monkey actively infected with HEV as a source of the agent. The method of
 XX the invention is used for detecting the presence of a viral agent in a
 XX sample of cultured cells infected with the agent and isolation of
 XX enterically transmitted nonA/nonB HEV agents or nucleic acid fragments
 XX produced by the agent. The present amino acid sequence represents the
 XX protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated
 XX from a Burmese strain of the hepatitis E virus.

Query Match 100.0% Score 4406; DB 22; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0.66 263;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 07 1 MKRPFLLLLMLFAPPGQSGRRGRSGSGGFWGDRVDSQPPALFYDPTN 60
 10 1 MKRPFLLLLMLFAPPGQSGRRGRSGSGGFWGDRVDSQPPALFYDPTN 60
 07 61 PRAPVTAAGAGPVRGAFARLGSAMWDQAGRTAVASRRPPTTAGAAPTAVAPADTP 120
 10 61 PRAPVTAAGAGPVRGAFARLGSAMWDQAGRTAVASRRPPTTAGAAPTAVAPADTP 120
 07 121 PWGDSRGLLRGYNSISGLSSVATGNTNVIYAAPLSPILPLQGTNHHMALEAS 180
 10 121 PWGDSRGLLRGYNSISGLSSVATGNTNVIYAAPLSPILPLQGTNHHMALEAS 180


```

121 PVPVDSGATIPQYNLSPLTSSVATGNTNVLVAAPLSPLLDGNTNTHMATEAS 190
131 NVAQVAPATIPYPIVPAVAGYATISEWPCITTTTPTSDVMNSITSDWPTVWPGI 240
141 NVAQVAPATIPYPIVPAVAGYATISEWPCITTTTPTSDVMNSITSDWPTVWPGI 240
241 ASFLVPSERLHYNQWRSVETSGVAFFATSCIOMIOTHSIVNSYNTIYTGALSL 300
241 ASFLVPSERLHYNQWRSVETSGVAFFATSCIOMIOTHSIVNSYNTIYTGALSL 300
301 DPALEPFRNLTGNTNTVPSYSSTAFRECEPACGTAFITTTAATREMKLYFTSNG 360
301 DPALEPFRNLTGNTNTVPSYSSTAFRECEPACGTAFITTTAATREMKLYFTSNG 360
361 VGEIGRGIALTFNADTILGGIPTFLISSAGGCLFYSPVVSANGPTVKIYTSVENAQ 420
361 VGEIGRGIALTFNADTILGGIPTFLISSAGGCLFYSPVVSANGPTVKIYTSVENAQ 420
421 QDKGIATPHDIDLGESVWVLIQYUNQHEQKPTSPAPSPPPSVLPKANDVLWLSAABY 480
421 QDKGIATPHDIDLGESVWVLIQYUNQHEQKPTSPAPSPPPSVLPKANDVLWLSAABY 480
481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLGDRPLSTTQQYSKTFVLP 540
481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLGDRPLSTTQQYSKTFVLP 540
541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVNAAGHRVAISPTYTSLGAGVVSISAVAV 600
541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVNAAGHRVAISPTYTSLGAGVVSISAVAV 600
601 IAPHSALALLEDTLQYPARARTPDGCPGCRPIGLAGCAPQSTVAELQPLKKVKVKTREL 660
601 IAPHSALALLEDTLQYPARARTPDGCPGCRPIGLAGCAPQSTVAELQPLKKVKVKTREL 660

```

RESULT 13

AA991814 standard: Protein, 660 AA.

AA991814:

26-NOV-1996 (first entry)

Hepatitis E virus strain SAP 55 OFF 2.

Hepatitis E virus: HFV, SAP-55 strain; enteric transmission;
structural region; antigen; detection; antibody; vaccine;
immunisation; infection.

Hepatitis E virus

W09610580-A2.

11-APR-1996

03-OCT-1995: 95W0-US13102.

03-OCT-1994: 94US-0316765.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Emerson SU, Purcell RH, Tsarev SA.

WPI, 1996-209320/21.

N-PSDB: AAT27394.

Isolated and purified hepatitis E virus strain SAP 55 DNA - encodes
PT antigenic protein useful in diagnosis, prophylaxis and treatment of
hepatitis E virus infection

Claim 6: Pages 13-15; 121pp; English.

The present sequence is the protein prod. of ORF 2 from the
hepatitis E virus (HEV) strain SAP 55, which was implicated in an
enterically transmitted non-A, non-B hepatitis in Pakistan. The
protein encoded by the structural region of the virus (i.e. ORF 2),
which is capable of forming HEV like particles, is useful for the
detection of HEV antibodies (prof. 145 or 146) in blood, plasma,
sera, cerebrospinal fluid, tissue, urine or pleural fluid, the
protein, and anti-HEV antibodies generated using the protein, can
also be used in vaccines for immunising an animal against HEV
infection. The protein is identified as a band of greater than
50 kDa following SDS PAGE of cell lysates of insect cells infected
with a HEV cDNA 2 cont. baculovirus. i.e. the claimed recombinant
expression vectors pPIC9-1779, -1780 and -1781.

Sequence 660 AA:

Query Match 99.5% Score 3491; DB 17; Length 660;
Percent Similarity 99.4%; Pos. No. 263;
Matches 656; Conservative 1; Mismatches 3; Indels 0; gaps 0;

```

QY 1 MRPRPILLILLMLPMLPAPPVQDPSRRRGKRSQSGSGGFWQVUSLPPFAIPYHPIN 60
DB 1 MRPRPILLILLMLPMLPAPPVQDPSRRRGKRSQSGSGGFWQVUSLPPFAIPYHPIN 60
QY 61 PFAPDVTAAAGAGPRVQRPAPRLSSAWKQQAUPPAVASRRPTTAGAAPTATVAFAHTP 120
DB 61 PFAPDVTAAAGAGPRVQRPAPRLSSAWKQQAUPPAVASRRPTTAGAAPTATVAFAHTP 120
QY 121 PVTNVSSEALIPRQVHLSLSLTSVATGTRNVIYAAAPLSLTLQWNTNTHMATEAS 180
DB 121 PVTNVSSEALIPRQVHLSLSLTSVATGTRNVIYAAAPLSLTLQWNTNTHMATEAS 180
QY 181 NVAGYKVARAIQYPIVPAVAGYATISEWPCITTTTPTSDVMNSITSDWPTVWPGI 240
DB 181 NVAGYKVARAIQYPIVPAVAGYATISEWPCITTTTPTSDVMNSITSDWPTVWPGI 240
QY 241 ASFLVPSERLHYNQWRSVETSGVAFFATSCIOMIOTHSIVNSYNTIYTGALSL 300
DB 241 ASFLVPSERLHYNQWRSVETSGVAFFATSCIOMIOTHSIVNSYNTIYTGALSL 300
QY 301 DPALEPFRNLTGNTNTVPSYSSTAFRECEPACGTAFITTTAATREMKLYFTSNG 360
DB 301 DPALEPFRNLTGNTNTVPSYSSTAFRECEPACGTAFITTTAATREMKLYFTSNG 360
QY 361 VGEIGRGIALTFNADTILGGIPTFLISSAGGCLFYSPVVSANGPTVKIYTSVENAQ 420
DB 361 VGEIGRGIALTFNADTILGGIPTFLISSAGGCLFYSPVVSANGPTVKIYTSVENAQ 420
QY 421 QDKGIATPHDIDLGESVWVLIQYUNQHEQKPTSPAPSPPPSVLPKANDVLWLSAABY 480
DB 421 QDKGIATPHDIDLGESVWVLIQYUNQHEQKPTSPAPSPPPSVLPKANDVLWLSAABY 480
QY 481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLGDRPLSTTQQYSKTFVLP 540
DB 481 DOSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLGDRPLSTTQQYSKTFVLP 540
QY 541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVNAAGHRVAISPTYTSLGAGVVSISAVAV 600
DB 541 LRGLSPWEAGTTKAGYPYNYNTTASDOLLVNAAGHRVAISPTYTSLGAGVVSISAVAV 600
QY 601 IAPHSALALLEDTLQYPARARTPDGCPGCRPIGLAGCAPQSTVAELQPLKKVKVKTREL 660
DB 601 IAPHSALALLEDTLQYPARARTPDGCPGCRPIGLAGCAPQSTVAELQPLKKVKVKTREL 660

```

RESULT 14

AA991814 standard: Protein, 660 AA.

AA991814:

11-JUN-1999 (first entry)

Db 61 PFAPDVTAAAGAGPVVPQAPPIQSAWPHQAGPFAASPPPTIAGAAPLTAVAPAHDTF 120
QY 121 PVPDYDSKAILRQYNI STSPITSSVATGNI VLVAAPLSPLPIQUGNTHIMATEAS 180
Db 121 PVPDYDSKAILRQYNI STSPITSSVATGNI VLVAAPLSPLPIQUGNTHIMATEAS 180
QY 181 NYAQRVARATIRYRVPVNAAGYALISFWQTTTTPTSVFMNSITSDVPILVQPGI 240
Db 181 NYAQRVARATIRYRVPVNAAGYALISFWQTTTTPTSVFMNSITSDVPILVQPGI 240
QY 241 ASELVIPSERIHYRNQWRSVETSGVAEEATSGLVMLCHGSLVNSYNTPTGALGL 300
Db 241 ASELVIPSERIHYRNQWRSVETSGVAEEATSGLVMLCHGSLVNSYNTPTGALGL 300
QY 301 DFALPIFPFNTIPGNTNTRVSPYSTAPHRIKPCADGTAEITTAATREMKDLYFTSTNG 360
Db 301 DFALPIFPFNTIPGNTNTRVSPYSTAPHRIKPCADGTAEITTAATREMKDLYFTSTNG 360
QY 361 VREIGRGIATLIFNIADTILAGIPTETISSAGGQIFYSPVVSANGPPTVKLYTSVENAG 420
Db 361 VREIGRGIATLIFNIADTILAGIPTETISSAGGQIFYSPVVSANGPPTVKLYTSVENAG 420
QY 421 QUKGTAIPHDINIGESPVVIGQYDNQHEQDRPIPSAPSPPPSVLPAANDVLMLSLTAAY 480
Db 421 QUKGTAIPHDINIGESPVVIGQYDNQHEQDRPIPSAPSPPPSVLPAANDVLMLSLTAAY 480
QY 481 DQSTYGSSTGVPVYSDSVTLNVNATGAQAVARSLDWTKVTLGGRPISTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGVPVYSDSVTLNVNATGAQAVARSLDWTKVTLGGRPISTIQOYSKTFVLP 540
QY 541 LRGLSPWEAGTTKAGYPYNYNTTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGLSPWEAGTTKAGYPYNYNTTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSAIIPGHTIDYPARHTFEDFPGCRPIGLQCCAFQSTVAELQELKMKVKIREL 660
Db 601 LAPHSAIIPGHTIDYPARHTFEDFPGCRPIGLQCCAFQSTVAELQELKMKVKIREL 660

Search completed: August 5, 2003, 09:27:03
Job time : 88 secs

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA: 100.00% Score, 4087 Indels, 0 Mismatches, 0 Gaps

APPLICATION NUMBER: 050799736 GATA

FILING DATE: 20 APRIL 1992

CLASSIFICATION: 435

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: US 822,335

FILING DATE: 17 JAN 1992

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: US 505,868

FILING DATE: 05 APRIL 1990

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: US 420,921

FILING DATE: 13 OCTOBER 1989

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: US 467,486

FILING DATE: 16 JUNE 1989

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: US 336,672

FILING DATE: 11 APRIL 1989

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: US 208,947

FILING DATE: 17 JUNE 1988

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: 34,875

REGISTRATION NUMBER: 4600-0093,46

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 424-0860

TELEFAX: (415) 424-0960

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 660 amino acids

TYPE: amino acid

TOPOLGY: Linear

MOLECULE TYPE: protein

HYDROPHILIC: No

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: GIB 2, BODMA, FUSION 9

US 97 870 997A 19

Query Match: 100.00% Score: 4087 Indels: 0 Gaps: 0
Best Local Similarity: 100.00% Ident. No.: 240-294
Matches: 660; Conservat. 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 MPPPELLELLLEMLPPLPAPPPQSGRRRRRRSGGGGFWGDRVDSOFFEALPYLPTN 60

1 MPPPELLELLLEMLPPLPAPPPQSGRRRRRRSGGGGFWGDRVDSOFFEALPYLPTN 60

61 PEALGVIAAGAGPRVPPAPPLPSAMRQQAQRPVAVASRRPTTAGAAPLTAVAPAHPTP 120

61 PEALGVIAAGAGPRVPPAPPLPSAMRQQAQRPVAVASRRPTTAGAAPLTAVAPAHPTP 120

121 PVPVADSGALLRRQPNLSISPLSSVAIGTINLVLYAAAPLPLPLPQAGNIHIMRAIRAS 180

121 PVPVADSGALLRRQPNLSISPLSSVAIGTINLVLYAAAPLPLPLPQAGNIHIMRAIRAS 180

121 PVPVADSGALLRRQPNLSISPLSSVAIGTINLVLYAAAPLPLPLPQAGNIHIMRAIRAS 180

141 NYAAYPVKARALIRKPLVENVAGVYALISFEPWCTITPTISVLMMNSTTSDVRILVOPGI 240

141 NYAAYPVKARALIRKPLVENVAGVYALISFEPWCTITPTISVLMMNSTTSDVRILVOPGI 240

241 ASELVTPSERLHYKNGQWRVSPETSGVAEEFATISGLVMLCHHGLSVNSYINIFYLAGLL 300

241 ASELVTPSERLHYKNGQWRVSPETSGVAEEFATISGLVMLCHHGLSVNSYINIFYLAGLL 300

401 DEALELEPRLIPQNTIRVSRYSSTAKHRLRCAKCTAKELITTAATREMKDLYFTSNG 460

401 DEALELEPRLIPQNTIRVSRYSSTAKHRLRCAKCTAKELITTAATREMKDLYFTSNG 460

61 VTELEPRLIPQNTIRVSRYSSTAKHRLRCAKCTAKELITTAATREMKDLYFTSNG 420

61 VTELEPRLIPQNTIRVSRYSSTAKHRLRCAKCTAKELITTAATREMKDLYFTSNG 420

421 OPERATING SYSTEM: PC-DOS/MS-DOS
421 CURRENT APPLICATION DATA: 100.00% Score, 4087 Indels, 0 Mismatches, 0 Gaps
421 APPLICATION NUMBER: 050799736 GATA
421 FILING DATE: 20 APRIL 1992
421 CLASSIFICATION: 435
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: US 822,335
421 FILING DATE: 17 JAN 1992
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: US 505,868
421 FILING DATE: 05 APRIL 1990
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: US 420,921
421 FILING DATE: 13 OCTOBER 1989
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: US 467,486
421 FILING DATE: 16 JUNE 1989
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: US 336,672
421 FILING DATE: 11 APRIL 1989
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: US 208,947
421 FILING DATE: 17 JUNE 1988
421 PREVIOUS APPLICATION DATA:
421 APPLICATION NUMBER: 34,875
421 REGISTRATION NUMBER: 4600-0093,46
421 TELECOMMUNICATION INFORMATION:
421 TELEPHONE: (415) 424-0860
421 TELEFAX: (415) 424-0960
421 INFORMATION FOR SEQ ID NO: 19:
421 SEQUENCE CHARACTERISTICS:
421 LENGTH: 660 amino acids
421 TYPE: amino acid
421 TOPOLGY: Linear
421 MOLECULE TYPE: protein
421 HYDROPHILIC: No
421 ORIGINAL SOURCE:
421 INDIVIDUAL ISOLATE: GIB 2, BODMA, FUSION 9
421 US 97 870 997A 19

RESULT 12

PCT-US95-14704-13

Sequence 13, Application PCT/US95-14704

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIBODIES AND

FILE OF INVENTION: 02/23/1995

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bellmont & Associates

STREET: P.O. Box 60653

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0860

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS DOS

SOFTWARE: Patent In Release #1.0, Version 1.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 13/704-13

CLASSIFICATION:

APPROPRIATE INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 34,875

REFERENCE/WORK NUMBER: 4600-0093,41

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 424-0860

TELEFAX: (415) 424-0960

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 660 amino acids

TYPE: amino acid

TOPOLGY: Linear

MOLECULE TYPE: protein

HYDROPHILIC: No

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Hepatitis E Virus (Bodma strain)

INDIVIDUAL ISOLATE: 004-2

PCT-US95-14704-13

Query Match: 100.00% Score: 4087 Indels: 0 Gaps: 0
Best Local Similarity: 100.00% Ident. No.: 240-294
Matches: 660; Conservat. 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 MPPPELLELLLEMLPPLPAPPPQSGRRRRRRSGGGGFWGDRVDSOFFEALPYLPTN 60

1 MPPPELLELLLEMLPPLPAPPPQSGRRRRRRSGGGGFWGDRVDSOFFEALPYLPTN 60

61 PEALGVIAAGAGPRVPPAPPLPSAMRQQAQRPVAVASRRPTTAGAAPLTAVAPAHPTP 120

61 PEALGVIAAGAGPRVPPAPPLPSAMRQQAQRPVAVASRRPTTAGAAPLTAVAPAHPTP 120

```

QY 121 EVWVDSGKALLEGANLSTLTSSVATCTNIVYAAVLSHLLLELLELLEKHIMATAS 180
DB 121 EVWVDSGKALLEGANLSTLTSSVATCTNIVYAAVLSHLLLELLELLEKHIMATAS 180
QY 141 NYAGSVARATITRYRIVNVAATYATFISFWMPTTITPTSVMSRITSTUWPKVQVQCI 240
DB 141 NYAGSVARATITRYRIVNVAATYATFISFWMPTTITPTSVMSRITSTUWPKVQVQCI 240
QY 241 ASPTVITSPHIFVNSQWPSVTSVAPITATSCVIMQIPIOSI VNSVTNPVNTCAVGL 300
DB 241 ASPTVITSPHIFVNSQWPSVTSVAPITATSCVIMQIPIOSI VNSVTNPVNTCAVGL 300
QY 301 DFALPTFEPMLTPTNTNTPSPYSSSTAPRPIEPFAGGTAFITTAATSEIMKLYPTSTNG 360
DB 301 DFALPTFEPMLTPTNTNTPSPYSSSTAPRPIEPFAGGTAFITTAATSEIMKLYPTSTNG 360
QY 341 VETFGRIATLTENLAITLQRIPIETISSAGQOI FYSPIVVSANCTPIVKLYSVPMQAO 420
DB 341 VETFGRIATLTENLAITLQRIPIETISSAGQOI FYSPIVVSANCTPIVKLYSVPMQAO 420
QY 421 QCKGTATPHITLGGESVWVQVYNGHFGERTFSCAFSEFVSVKANIVWLISIAAFY 480
DB 421 QCKGTATPHITLGGESVWVQVYNGHFGERTFSCAFSEFVSVKANIVWLISIAAFY 480
QY 481 DQSTYRSTGKVVSQSVILVNVAATAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
DB 481 DQSTYRSTGKVVSQSVILVNVAATAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
QY 541 LRKLSIWPACTKAGYPYNYNTASDQILVENAACHRVATISYTTSLGAGVTSISAVAV 600
DB 541 LRKLSIWPACTKAGYPYNYNTASDQILVENAACHRVATISYTTSLGAGVTSISAVAV 600
QY 601 LAPHSAIATLTETIYAPARHITDRTFPRCPPIGLQSTAPSTVAPLQPIKMKVSKDPEI 660
DB 601 LAPHSAIATLTETIYAPARHITDRTFPRCPPIGLQSTAPSTVAPLQPIKMKVSKDPEI 660

RESULT 13
US-09-172-699-2
: Sequence 2 Application US/09172699A
: Patent No. 6514690
: GENERAL INFORMATION:
: APPLICANT: Anderson, David A.
: APPLICANT: Locarnini, Stephen A.
: APPLICANT: Torressi, Joseph
: APPLICANT: Hui, Zhuang
: APPLICANT: Li, Fan
: TITLE OF INVENTION: IMMUNOREACTIVE ANTIBODIES OF HEPATITIS E VIRUS
: FILE REFERENCE: Davies Col. C, 96
: CURRENT APPLICATION NUMBER: US/09/172,699A
: CURRENT FILING DATE: 1998-10-14
: EARLIER APPLICATION NUMBER: 02/013,132
: EARLIER FILING DATE: 1996-06-20
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2
: LENGTH: 660
: TYPE: PRI
: ORGANISM: Homo sapiens
US-09-172-699-2

```

```

Query Match 90.7% Score 3499 DB 4: Length 660
Best Local Similarity 90.7% Prod. No. 1,50,293
Matches 658, Conservative 1, Mismatches 1, Indels 0, Gaps 0

```

```

QY 1 MRPRPTILLMLPMLPAPRPGSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 60
DB 1 MRPRPTILLMLPMLPAPRPGSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 60
QY 61 PRAPVTAAATATFETATFETATFETATFETATFETATFETATFETATFETATFETATFET 120
DB 61 PRAPVTAAATATFETATFETATFETATFETATFETATFETATFETATFETATFETATFET 120

```

```

QY 121 EVWVDSGKALLEGANLSTLTSSVATCTNIVYAAVLSHLLLELLELLEKHIMATAS 180
DB 121 EVWVDSGKALLEGANLSTLTSSVATCTNIVYAAVLSHLLLELLELLEKHIMATAS 180
QY 141 NYAGSVARATITRYRIVNVAATYATFISFWMPTTITPTSVMSRITSTUWPKVQVQCI 240
DB 141 NYAGSVARATITRYRIVNVAATYATFISFWMPTTITPTSVMSRITSTUWPKVQVQCI 240
QY 241 ASPTVITSPHIFVNSQWPSVTSVAPITATSCVIMQIPIOSI VNSVTNPVNTCAVGL 300
DB 241 ASPTVITSPHIFVNSQWPSVTSVAPITATSCVIMQIPIOSI VNSVTNPVNTCAVGL 300
QY 301 DFALPTFEPMLTPTNTNTPSPYSSSTAPRPIEPFAGGTAFITTAATSEIMKLYPTSTNG 360
DB 301 DFALPTFEPMLTPTNTNTPSPYSSSTAPRPIEPFAGGTAFITTAATSEIMKLYPTSTNG 360
QY 341 VETFGRIATLTENLAITLQRIPIETISSAGQOI FYSPIVVSANCTPIVKLYSVPMQAO 420
DB 341 VETFGRIATLTENLAITLQRIPIETISSAGQOI FYSPIVVSANCTPIVKLYSVPMQAO 420
QY 421 QCKGTATPHITLGGESVWVQVYNGHFGERTFSCAFSEFVSVKANIVWLISIAAFY 480
DB 421 QCKGTATPHITLGGESVWVQVYNGHFGERTFSCAFSEFVSVKANIVWLISIAAFY 480
QY 481 DQSTYRSTGKVVSQSVILVNVAATAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
DB 481 DQSTYRSTGKVVSQSVILVNVAATAGAVAKSLIWKVILGKPLSLIAYSKIFVLP 540
QY 541 LRKLSIWPACTKAGYPYNYNTASDQILVENAACHRVATISYTTSLGAGVTSISAVAV 600
DB 541 LRKLSIWPACTKAGYPYNYNTASDQILVENAACHRVATISYTTSLGAGVTSISAVAV 600
QY 601 LAPHSAIATLTETIYAPARHITDRTFPRCPPIGLQSTAPSTVAPLQPIKMKVSKDPEI 660
DB 601 LAPHSAIATLTETIYAPARHITDRTFPRCPPIGLQSTAPSTVAPLQPIKMKVSKDPEI 660

RESULT 14
US-09-462-606-12
: Sequence 12 Application US/090942606
: Patent No. 6443408
: GENERAL INFORMATION:
: APPLICANT: MENG, XIANG-JIN
: APPLICANT: Emerson, Suzanne H.
: APPLICANT: Purcell, Robert H.
: TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
: FILE REFERENCE: 20240267051
: CURRENT FILING DATE: 2000-06-12
: PRIOR APPLICATION NUMBER: US 60/053069
: PRIOR FILING DATE: 1997-07-18
: EARLIER APPLICATION NUMBER: 02/013,132
: EARLIER FILING DATE: 1996-06-20
: NUMBER OF SEQ ID NOS: 65
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 12
: LENGTH: 660
: TYPE: PRI
: ORGANISM: Hepatitis E virus
US-09-462-606-12

```

```

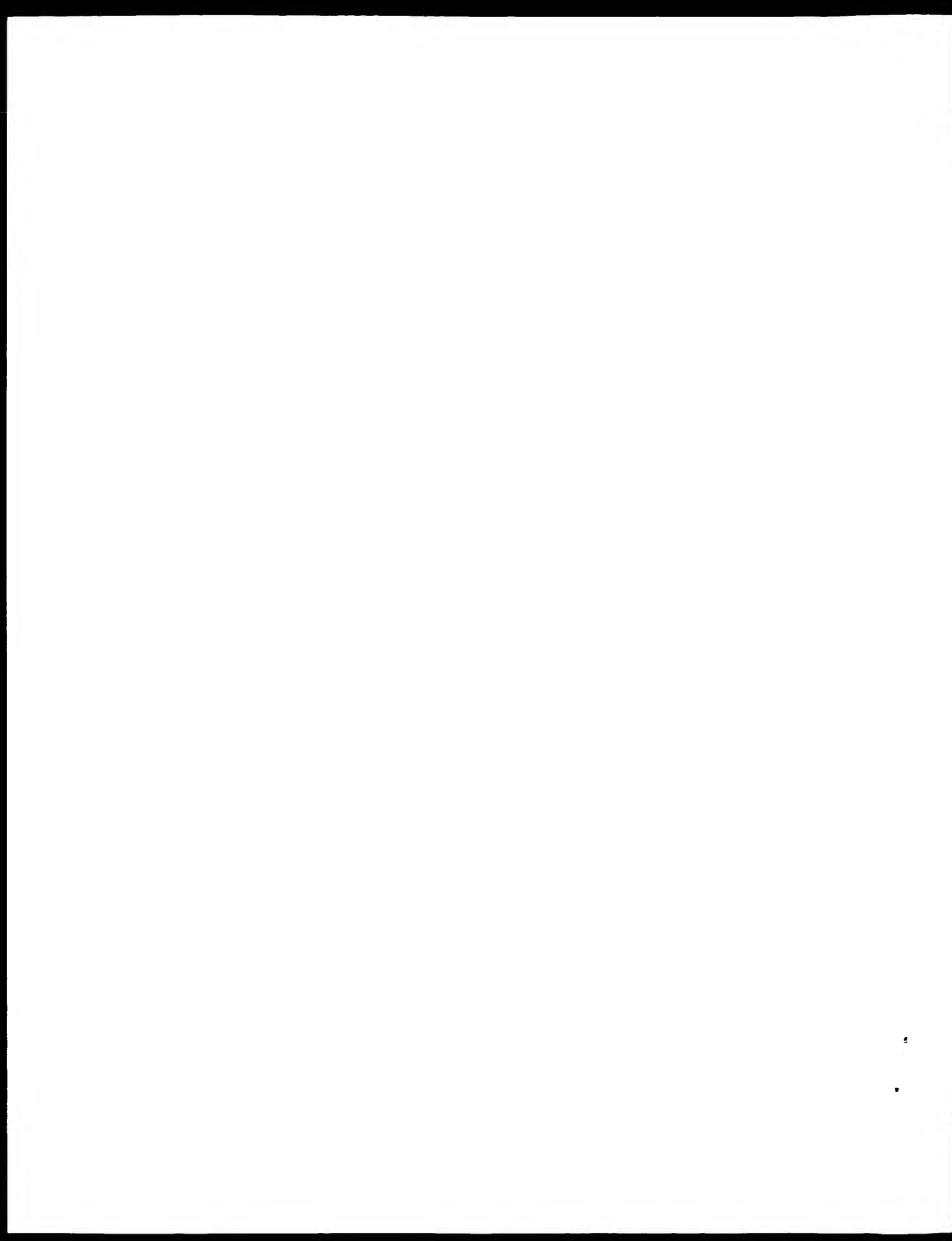
Query Match 99.5% Score 4991 DB 4: Length 660
Best Local Similarity 99.5% Prod. No. 7,70,204
Matches 676, Conservative 1, Mismatches 0, Indels 0, Gaps 0

```

```

QY 1 MRPRPTILLMLPMLPAPRPGSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 60
DB 1 MRPRPTILLMLPMLPAPRPGSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 60
QY 61 PRAPVTAAATATFETATFETATFETATFETATFETATFETATFETATFETATFETATFET 120
DB 61 PRAPVTAAATATFETATFETATFETATFETATFETATFETATFETATFETATFETATFET 120

```


```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 660 amino acids
: TYPE: amino acid
: STRANDEDNESS: Hepatitis E Virus (Burma strain)
: ORF-2
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-769-066-13

Query Match: 130.0% Score 3408, DB 11, Length 660,
Best Local Similarity 100.0% Pred. No. 3,26-276;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPPPTIIIMETPMIAPHPGSPSPHPKSPSSGGSGFWGLPVLSTPEALPYIPPTN 60
DB 1 MRPPPTIIIMETPMIAPHPGSPSPHPKSPSSGGSGFWGLPVLSTPEALPYIPPTN 60

QY 61 PPAPIVTAACAGPRVHPQAPPLGSAWPDQAGPAPVAPSPPPPTTAGAAPTAVAPADPP 120
DB 61 PPAPIVTAACAGPRVHPQAPPLGSAWPDQAGPAPVAPSPPPPTTAGAAPTAVAPADPP 120

QY 121 PVPQWISRGATIPQYNIESTLSSVATGNI.VLYAAFLSPCLPQDGTNTHIMATEAS 180
DB 121 PVPQWISRGATIPQYNIESTLSSVATGNI.VLYAAFLSPCLPQDGTNTHIMATEAS 180

QY 181 NYAQYVAPATIPYFTYVPAAGGYAIEISEWQTITITISVNMNSIISTEPVLVFP31 240
DB 181 NYAQYVAPATIPYFTYVPAAGGYAIEISEWQTITITISVNMNSIISTEPVLVFP31 240

QY 241 ASFTVIPSFTHYNQWPSVFTSCVAFPAATSGI.VMGLHRSIVRSYNTNTVYTAAG31 300
DB 241 ASFTVIPSFTHYNQWPSVFTSCVAFPAATSGI.VMGLHRSIVRSYNTNTVYTAAG31 300

QY 301 DPALEIPERNITPGNTNTRVSRYSSTAPHPLRPADGTARLTATTAATPEMKDLYFTSNG 360
DB 301 DPALEIPERNITPGNTNTRVSRYSSTAPHPLRPADGTARLTATTAATPEMKDLYFTSNG 360

QY 361 VETGPGIALTFLENALTLLGGPLTELISACGQLEYSPRVVSANGEPVKLYTSVENAQ 420
DB 361 VETGPGIALTFLENALTLLGGPLTELISACGQLEYSPRVVSANGEPVKLYTSVENAQ 420

QY 421 QKKGIAIPHIDMGESPVVTDQYNQHQHQHPPTSPAPSPFVSLEANDVLWLSITAAEY 480
DB 421 QKKGIAIPHIDMGESPVVTDQYNQHQHQHPPTSPAPSPFVSLEANDVLWLSITAAEY 480

QY 481 DQSTGSGTGPVYVSDSVTLVNVATGACAVAPSI.PWTKVTLDGPPSLTQQYSKTFEVL 540
DB 481 DQSTGSGTGPVYVSDSVTLVNVATGACAVAPSI.PWTKVTLDGPPSLTQQYSKTFEVL 540

QY 541 LRKGLSPWFACITKAGYPYNYNTASDQILVFNAAQHPVAISYVYTSICAGHPVSI.SAVAV 600
DB 541 LRKGLSPWFACITKAGYPYNYNTASDQILVFNAAQHPVAISYVYTSICAGHPVSI.SAVAV 600

QY 601 LAPHSAIAGLEDTLQYPAFAHTFQGFPEFEPGLSGPQAPQSTVAEGQPLFMEVSKTPE 660
DB 601 LAPHSAIAGLEDTLQYPAFAHTFQGFPEFEPGLSGPQAPQSTVAEGQPLFMEVSKTPE 660

RESULT 2
US-09-851-410-a
: Sequence B, Application US/09851410
: Publication No. us09040124510a1
: GENERAL INFORMATION:
: APPLICANT: Reyes, Greider, E
: Yarbough, Patrice O
: Bradley, Daniel W
: Krawczynski, Krzysztof Z
: Tam, Albert
: Fry, Kirk E
: TITLE OF INVENTION: DNA Sequences of Putatively Transmitted

```

```

: NUMBER OF SEQUENCES: 20
: CORRESPONDENT ADDRESS:
: ADDRESSEE: TechTarget & Associates
: STREET: 450 Cambridge Avenue, Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: CREATING SYSTEM: PC 4.5 MS-DOS
: SOFTWARE: PatentID Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/851,410
: FILING DATE: 07 MAY 2001
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/129,275
: FILING DATE: UNKNOWN
: APPLICATION NUMBER: 08 07/681,078
: FILING DATE: 05 APR 1991
: APPLICATION NUMBER: 08 07/505,886
: FILING DATE: 05 APR 1998
: APPLICATION NUMBER: 08 07/420,921
: FILING DATE: 13 OCT 1989
: APPLICATION NUMBER: 08 07/367,486
: FILING DATE: 16 JUN 1989
: APPLICATION NUMBER: 08 07/304,672
: FILING DATE: 11 APR 1989
: APPLICATION NUMBER: 08 07/298,992
: FILING DATE: 17 JUN 1988
: ACTING AGENT INFORMATION:
: NAME: Petalibody, Joanne R
: REGISTRATION NUMBER: 42,995
: REFERENCE NUMBER: 4,689,014-24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 324-0880
: TELEFAX: (650) 324-0960
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 660 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-a

Query Match: 100.0% Score 3408, DB 11, Length 660;
Best Local Similarity 100.0% Pred. No. 3,26-276;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPPPTIIIMETPMIAPHPGSPSPHPKSPSSGGSGFWGLPVLSTPEALPYIPPTN 60
DB 1 MRPPPTIIIMETPMIAPHPGSPSPHPKSPSSGGSGFWGLPVLSTPEALPYIPPTN 60

QY 61 PPAPIVTAACAGPRVHPQAPPLGSAWPDQAGPAPVAPSPPPPTTAGAAPTAVAPADPP 120
DB 61 PPAPIVTAACAGPRVHPQAPPLGSAWPDQAGPAPVAPSPPPPTTAGAAPTAVAPADPP 120

QY 121 PVPQWISRGATIPQYNIESTLSSVATGNI.VLYAAFLSPCLPQDGTNTHIMATEAS 180
DB 121 PVPQWISRGATIPQYNIESTLSSVATGNI.VLYAAFLSPCLPQDGTNTHIMATEAS 180

QY 181 NYAQYVAPATIPYFTYVPAAGGYAIEISEWQTITITISVNMNSIISTEPVLVFP31 240
DB 181 NYAQYVAPATIPYFTYVPAAGGYAIEISEWQTITITISVNMNSIISTEPVLVFP31 240

QY 241 ASFTVIPSFTHYNQWPSVFTSCVAFPAATSGI.VMGLHRSIVRSYNTNTVYTAAG31 300
DB 241 ASFTVIPSFTHYNQWPSVFTSCVAFPAATSGI.VMGLHRSIVRSYNTNTVYTAAG31 300

QY 301 DPALEIPERNITPGNTNTRVSRYSSTAPHPLRPADGTARLTATTAATPEMKDLYFTSNG 360
DB 301 DPALEIPERNITPGNTNTRVSRYSSTAPHPLRPADGTARLTATTAATPEMKDLYFTSNG 360

QY 361 VETGPGIALTFLENALTLLGGPLTELISACGQLEYSPRVVSANGEPVKLYTSVENAQ 420
DB 361 VETGPGIALTFLENALTLLGGPLTELISACGQLEYSPRVVSANGEPVKLYTSVENAQ 420

QY 421 QKKGIAIPHIDMGESPVVTDQYNQHQHQHPPTSPAPSPFVSLEANDVLWLSITAAEY 480
DB 421 QKKGIAIPHIDMGESPVVTDQYNQHQHQHPPTSPAPSPFVSLEANDVLWLSITAAEY 480

QY 481 DQSTGSGTGPVYVSDSVTLVNVATGACAVAPSI.PWTKVTLDGPPSLTQQYSKTFEVL 540
DB 481 DQSTGSGTGPVYVSDSVTLVNVATGACAVAPSI.PWTKVTLDGPPSLTQQYSKTFEVL 540

QY 541 LRKGLSPWFACITKAGYPYNYNTASDQILVFNAAQHPVAISYVYTSICAGHPVSI.SAVAV 600
DB 541 LRKGLSPWFACITKAGYPYNYNTASDQILVFNAAQHPVAISYVYTSICAGHPVSI.SAVAV 600

QY 601 LAPHSAIAGLEDTLQYPAFAHTFQGFPEFEPGLSGPQAPQSTVAEGQPLFMEVSKTPE 660
DB 601 LAPHSAIAGLEDTLQYPAFAHTFQGFPEFEPGLSGPQAPQSTVAEGQPLFMEVSKTPE 660

```



```

1 Yarlough, Patricia G.
2 Zhong, Yilan
3
4 TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREOF
5 NUMBER OF SEQUENCES: 41
6 REFERENCE ADDRESS
7 ADDRESS: Dohlinet & Associates
8 STREET: 360 Cambridge Ave., Suite 250
9 CITY: Palo Alto
10 STATE: CA
11 COUNTRY: USA
12 ZIP: 94306
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patient In Release #1.0, Version #1.25
19
20 APPLICATION NUMBER: US2002/010746A1
21 FILING DATE: 24 Jan 2001
22 CLASSIFICATION: C01K 3/00
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 08/542,644
26 FILING DATE: Unknown
27
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Fadhil, Gary E.
30 REGISTRATION NUMBER: 34,875
31 REFERENCE/AGENT NUMBER: 4500-0294, 40
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: (415) 324-0880
34 TELEFAX: (415) 324-0960
35
36 INFORMATION FOR SEQ ID NO: 15:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 549 amino acids
39 TYPE: amino acid
40 STRAND: SENSE
41
42 Description: Hepatitis E virus (Burma strain)
43
44 Topology: linear
45
46 MOLECULE TYPE: protein
47 HYDROPHobic: No
48 ORIGIN: SOURCE:
49 SEQUENCE DESCRIPTION: SEQ ID NO: 15:
50
51 us-09-769-066-25
52
53 Query Match: 62.5% Score: 2814; DB 10; Length: 549;
54 Best Local Similarity: 100.0%; Ident: 100.0%; Mismatch: 0; Gaps: 0;
55 Matches: 549; Conserved: 0; Mismatch: 0; Labels: 0;
56
57 112 AVAAHAPPPVWDISRAILRRYNLSISPLSSSVATGNIWYAAPISDPIPLQDGIN 171
58 1 AVAAHAPPPVWDISRAILRRYNLSISPLSSSVATGNIWYAAPISDPIPLQDGIN 69
59
60 174 TRMATAHARYAGTAVAKATKVLVNAVAYALSTFWQTTPTSVLMNSTSTG 281
61 1 TRMATAHARYAGTAVAKATKVLVNAVAYALSTFWQTTPTSVLMNSTSTG 120
62
63 244 VRLVYVGLHASELVIFSRKGRFRQWRSVELGVAVELASGLVMALHGSVLNYSIN 291
64 1 VRLVYVGLHASELVIFSRKGRFRQWRSVELGVAVELASGLVMALHGSVLNYSIN 190
65
66 292 DYKAGLLDEAFLEERNEIPGNTNTRVSKYSTAHKLRKGAETAEIILIAIRPKK 351
67 1 DYKAGLLDEAFLEERNEIPGNTNTRVSKYSTAHKLRKGAETAEIILIAIRPKK 250
68
69 352 IATYSTNGVIGRGIALIFNLADTLGRLTELISSAGGQFYSRPVVSANGPEPTVK 411
70 1 IATYSTNGVIGRGIALIFNLADTLGRLTELISSAGGQFYSRPVVSANGPEPTVK 300
71
72 412 IYTSVENAQQKQGLAIFHIDHESPPVLTPTNDRQGRPTSPAPSPFSPVLRANVL 471
73 1 IYTSVENAQQKQGLAIFHIDHESPPVLTPTNDRQGRPTSPAPSPFSPVLRANVL 350
74
75 472 WLSLTAALYDQSTSTSLGVVGGVLYNVAAGAAVAFSLDILKVTLLGKPSLLG 541
76 1 WLSLTAALYDQSTSTSLGVVGGVLYNVAAGAAVAFSLDILKVTLLGKPSLLG 400
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```


1 SOFTWARE: Patent In Release #1.0, Version #1.25
 2 CURRENT APPLICATION DATA:
 3 APPLICATION NUMBER: US 09/066-26
 4 FILING DATE: 24 Jan 2001
 5 CLASSIFICATION: Unknown
 6 PRIOR APPLICATION DATA:
 7 APPLICATION NUMBER: Unknown
 8 FILING DATE: Unknown
 9 ATTORNEY/AGENT INFORMATION:
 10 NAME: Fabian, Gary R.
 11 REGISTRATION NUMBER: 44,875
 12 REFERENCE/PROCKET NUMBER: 4600 0294.40
 13 TELECOMMUNICATION INFORMATION:
 14 TELEPHONE: (415) 424-0680
 15 TELEFAX: (415) 424-0960
 16 INFORMATION FOR SEQ ID NO: 16:
 17 SEQUENCE CHARACTERISTICS:
 18 LENGTH: 549 amino acids
 19 TYPE: amino acid
 20 STRANDEDNESS: Linear
 21 MOLECULAR TYPE: Protein
 22 HYPOTHEICAL: No
 23 ORIGINAL SOURCE:
 24 SEQUENCE DESCRIPTION: SEQ ID NO: 16:
 25

26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500

501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

Search completed: August 5, 2003, 09:31:42
J.E. file: 27 seqs

```
QY 511 APSLWTFVLDGRLSTGCGYSKIEFVLHPELSFWAGTTEACYPYNYPTASDQL 570
DB 302 ARSLTWSKVTLDGRLPTTIOQYSKTFYVLPURKLSFWAGTTKAGYPYNYTTASDQL 361
QY 571 VNAAGHRVAISTYTSISLACPVSSAVAVLAPHSALALEDTLDYPARAHFTDDPCPR 630
DB 462 TENAAGHRVAISTYTSISLACPVSSAVAVLAPHSALALEDTLDYPARAHFTDDPCPR 421
QY 641 RPLSLQGCAGFQSTVAELQRLMKVKGKTR 659
DB 422 RTLSLQGCAGFQSTVAELQRLMKVKGKTR 450
```

RESULT 15

```
US-09-468-147-199
: Sequence 199, Application US/09468147A
: Publication No. US20030049601A1
: GENERAL INFORMATION:
: APPLICANT: Abbott Laboratories
: APPLICANT: Schlaudor, George G.
: APPLICANT: Erker, James C.
: APPLICANT: Desai, Suresh M.
: APPLICANT: Dawson, George J.
: APPLICANT: Mushahwar, I. K.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
: TITLE OF INVENTION: HEPATITIS E VIRUS
: FILE REFERENCE: 6252 US.P1
: CURRENT APPLICATION NUMBER: US/09/468,147A
: CURRENT FILING DATE: 1999-12-21
: EARLIER APPLICATION NUMBER: US 09/173,141
: EARLIER FILING DATE: 1998-10-15
: EARLIER APPLICATION NUMBER: US 60/061,199
: EARLIER FILING DATE: 1997-10-15
: NUMBER OF SEQ ID NOS: 258
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 199
: LENGTH: 338
: TYPE: PRT
: ORGANISM: Hepatitis E Virus
: FEATURE:
: OTHER INFORMATION: cksort2m-2.pep
US-09-468-147-199
```

```
Query Match 47.0% Score 1601; DB 11; Length 338;
Best Local Similarity 93.6% Pred. No. 1.4e-125;
Matches 305; Conservative 14; Mismatches 7; Indels 0; Gaps 0;

QY 334 GAKGTAEITTTAATREMKDLYFTSTNGVGHIGRGIALTFNLNVLGLLGLPTELISSAG 393
DB 4 GAKGTAEITTTAATREMKDLHFAGTNGVGVSRGIALTFNLNVLGLLGLPTELISSAG 63
QY 454 QLEYSRPVVSANGEPYVLYTSVENAGGKGIAPHIDLGESFVVLQYVWVNHQEDRPT 453
DB 64 QIFYSRPVVSANGEPYVLYTSVENAGGKGIAPHIDLGESFVVLQYVWVNHQEDRPT 123
QY 454 PSPAPSPFVSRLANDVWLSTAAEYDQSTYGSSTGCVVSVLNVAVATCAQAVARS 513
DB 124 PSPAPSPFVSRLANDVWLSTAAEYDQSTYGSSTGCVVSVLNVAVATCAQAVARS 183
QY 514 LQWTKVTLNRPPISTIOQYSKTFYVLPURKLSFWAGTTKAGYPYNYTTASDQLN 573
DB 184 LDWSKVTLDGRLPTTIOQYSKTFYVLPURKLSFWAGTTKAGYPYNYTTASDQLN 243
QY 574 AAGHRVAISTYTSISLACPVSSAVAVLAPHSALALEDTLDYPARAHFTDDPCPR 633
DB 244 AAGHRVAISTYTSISLACPVSSAVAVLAPHSALALEDTLDYPARAHFTDDPCPR 403
QY 634 GLQSCAFQSTVAELQRLMKVKGKTR 659
DB 404 GLQSCAFQSTVAELQRLMKVKGKTR 329
```


Db 589 GSUTLTAATL-----DTTFTVSGSGPGLMMDAALTAHAPGLTSLP 630
 QY 402 FAIFIFPNTTPGNTNTPVSPYSTAP-----HPLPPGA----- 336
 Db 631 -----TLLTGTHTDLSHATTAAQPLRHGMITTAQCHLTAYGTHTIQIAPPLD 682
 QY 336 -----DGAELFTTTAATRPKKDLYPSTNVRGIGRCIALTIENLADTLGGIPPEI 387
 Db 683 NTGRTIATNTLIDHTAALENTGTEHSTAT-----GNPFTTITTEINEAGHLEENPTEIL 739
 QY 488 I-----SSAGGOLFYSRPVV-----SANCEPTVKLYTSVENAQQKGIAPIH 429
 Db 740 TIGTWTNTGSHQITGQATHAHTIDNPNGTHTATGSPDLDTTEILNN--GUNGELTTT 797
 QY 430 D-IDLGSRVVOQYDNQH-FQDRPTPSAP-----SPFESVIPANTVWI----- 473
 Db 798 DALTLTAASLL-----NORCTTHAAGPAHETLTCGLTNTAGSLAQIAHLIMLISALINKS 852
 QY 474 -SLTAAEYDOSTYSS-----TGPV-----YVSDSVIL 500
 Db 853 GTLASGLTLDGTCHLIDNTGPGGTTHGNLILHSGLPLNTAGLQATAAALTDTGAATL 912
 QY 501 VNVATGACAVARSIDKVTIDGRPTSTQQYSKTFEVLPLGRKISFWPAGTKACAPYN 560
 Db 913 INRDGGALLAATLDTALTLDNRG-CTIDSQAT-----HLHTTILDNTTAGHSS 963
 QY 561 YNTTASQQLIVENAAG 576
 Db 964 SCTIQLDGTTLTNTGG 979

RESULT 10

hemagglutinin-like secreted protein XP2775 [imported] Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20 Aug 2000
 C:Accession: B82519
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82519
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 33455-SSIM-
 A:Class references: GR A1004987; GR A5001849; NID 93108063; FIIN AAF65601; ASPHA-30001
 A:Experimental source: strain 9a5c
 R:Rimpon, A I G.; Rainach, F C.; Arruda, P.; Abreu, F A.; Accencio, M.; Alvarado, R.; B.
 Briones, M E S.; Buena, M F F.; Camargo, A A.; Canarao, L E A.; Carriao, D M.; Gallet, R.
 as Neto, E.; Pacheco, C.; El-toumy, H.; Fariacouli, A P.; Ferreira, A J S.
 submitted to GenBank, June 2003
 A:Authors: Ferreira, V C A.; Farias, J A.; Farias, J S.; Farias, S C.; Farias, M C.; Farias,
 J D.; Junqueira, M L.; Komper, E L.; Kitajima, J P.; Krueger, J E.; Kuramao, E E.; Lages,
 chado, M A.; Madeira, A B N.; Madeira, H M F.; Marino, C L.; Marques, M V.; Martins, E.
 A:Authors: Martins, E M F.; Matsukuma, A Y.; Meuck, C F M.; Miracca, E C.; Miyaki, C Y.;
 F. G.; Nunes, L R.; Oliveira, M A.; de Oliveira, M C.; de Oliveira, R C.; Palmieri, D A.;
 Rodrigues, V.; Rosa, A I.; de M. de Rosa, F. V. E.; de Sa, P G.; Santelli, R V.; Sawasak
 A:Authors: da Silva, A C P.; da Silva, F R.; da Silva, A M.; Silva Jr., W A.; da Silva,
 M.; Isibako, M H.; Vallada, H.; Van Sluys, M A.; Venceski, Almeida, S.; Venceski, A.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF2775

Query Match 4.9% Score 135.5 DB 2 Length 3455
 Best Local Similarity 21.9% Pred. No. 8;
 Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27.
 QY 97 ASREFTTAAAEELTAVAFADHTP-----FVP-----QVSGSAIL 122
 Db 429 ASPADPT-----PVASVPPATTPPSTGMPISPHQATAPVLAAGRLTLENIDNPCH 483

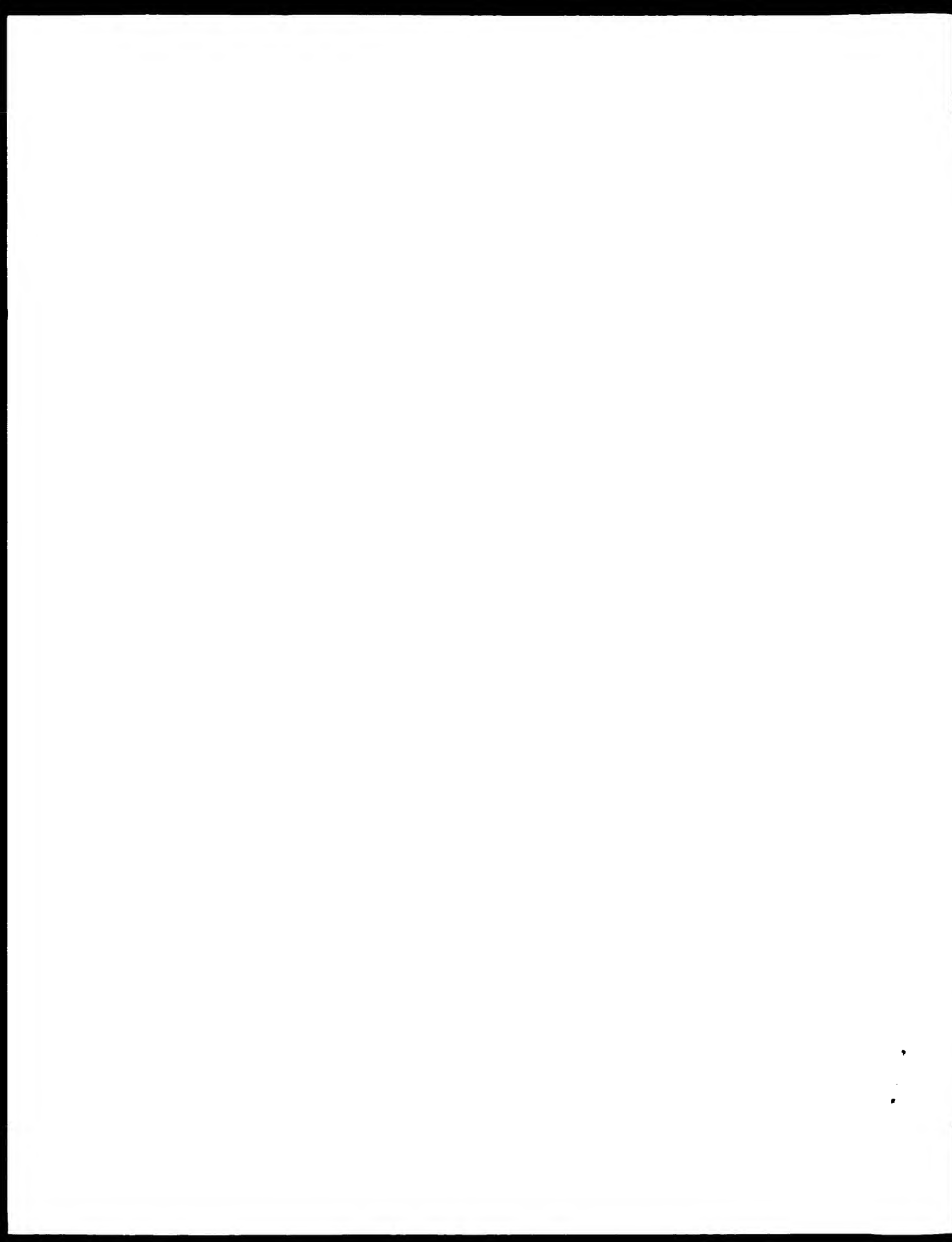
QY 122 KPOYNLTSPRLSSVAQVQNLNLYAALPDLPLPFGQCTNTH-----IMAFEA-----SN 181
 Db 484 TAGGATDAVLETTLDNPNV-----AALNPLTGGGQPLNAGCTTTATFATIPPTNN 637
 QY 182 YAOYKVARATIPYEPVPPNAGQYALISLSPWQTHITPISVIUMNSLSTVWHLVLP3A 241
 Db 538 AACQI PVNRTIDTAQHI NNASGHLITP3--POSAPHTIANLNTSTL-----ASA 689
 QY 242 SEGVTLDELHYRQWESVLTGVALBLAELERVWHTERLVRYNTPTPYGASLH 401
 Db 589 GSUTLTAATL-----DTTGTGSGGQGLMDAALTAHAPGLTSLP 630
 QY 302 FALDELPPNLTGNTNTPVSPYSTAP-----HRLRGA----- 445
 Db 631 -----TLLTGTHTDLSHATTAAQPLRHGMITTAQCHLTAYGTHTIQIAPPLD 682
 QY 336 -----DGAELFTTTAATRPKKDLYPSTNVRGIGRCIALTIENLADTLGGIPPEI 387
 Db 683 NTGRTIATNTLIDHTAALENTGTEHSTAT-----GNPFTTITTEINEAGHLEENPTEIL 739
 QY 488 I-----SSAGGOLFYSRPVV-----SANCEPTVKLYTSVENAQQKGIAPIH 429
 Db 798 DALTLTAASLL-----NORCTTHAAGPAHETLTCGLTNTAGSLAQIAHLIMLISALINKS 852
 QY 474 -SLTAAEYDOSTYSS-----TGPV-----YVSDSVIL 500
 Db 853 GTLASGLTLDGTCHLIDNTGPGGTTHGNLILHSGLPLNTAGLQATAAALTDTGAATL 912
 QY 501 VNVATGACAVARSIDKVTIDGRPTSTQQYSKTFEVLPLGRKISFWPAGTKACAPYN 560
 Db 913 INRDGGALLAATLDTALTLDNRG-CTIDSQAT-----HLHTTILDNTTAGHSS 963
 QY 561 YNTTASQQLIVENAAG 576
 Db 964 SCTIQLDGTTLTNTGG 979

RESULT 11

T21460
 hypothetical protein EK945.10 Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21460
 R:Wilkinson, J.
 submitted to the EMBL Data Library, March 1995
 A:Reference number: 219425
 A:Accession: T21460
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1796-WIL-
 A:Class references: IMBL348952; FIIN CAA8649.1; GSPB C83022; GESP EK945.10
 A:Experimental source: Glove F27E5
 C:Genetics:
 A:Gene: GESP:EK945.10
 A:Map position: 2
 A:Contents: annotation

Query Match 3.9% Score 134; DB 2 Length 766;
 Best Local Similarity 20.6% Pred. No. 1.4;
 Matches 121; Conservative 61; Mismatches 208; Indels 176; Gaps 24;

QY 63 APVTAAACACQVPPQPAHPSAHPGSAHPGACAGANVASPPPPHSAADPAAAPAPPLPV 122
 Db 242 ARQEMELASPIKEMESLAGLNLYLIESISLTPTPTTIVTTTIVTIV 296
 QY 123 TAVEMGACALRYALITLTSVATQTMVLYAFSLTLLTSGGNHIMAHAGNY 182
 Db 297 P-----TSTCTVTAMSTSS-----TGTSTTSTSTSTSTSTST 329



GenInfo version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

CM protein protein search, using sw model

Run on: August 5, 2003, 09:17:48 ; Search time 27 Seconds
(without alignments)
1149,542 Million cell updates/sec

Library: us-09-851-410a-8
Database: 3406
Sequence: 1 MRPRLILLMLLMFLMLAP.....QSTVALQLRLKMKVGTREL 660

Scoring table: BSUM62
Gapop 10.0 ; Gapext 0.5

Searches: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum db seq length: 0
Maximum db seq length: 2000000000

Post processor: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

red. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6408	100.0	660	1 VST2_HEVR0	P29326 hepatitis e
2	5491	99.5	660	1 VST2_HEVR0	P33426 hepatitis e
3	3467	98.8	660	1 VST2_HEVR0	Q04611 hepatitis e
4	3144.5	93.7	659	1 VST2_HEVR0	Q03500 hepatitis e
5	2446	71.8	485	1 VST2_HEVR0	Q06270 hepatitis e
6	143	4.2	1041	1 ECT2_YEAST	P42835 saccharomye
7	137.5	4.0	1083	1 T2E3_HUMAN	Q6C269 homo sapien
8	137	4.0	5793	1 M05H_HUMAN	Q06284 homo sapien
9	133	3.9	670	1 YF03_SCHPO	Q13854 schizosacch
10	133	3.9	1802	1 HRP1_YEAST	P41809 saccharomye
11	133	3.9	3178	1 YS89_CAEEL	Q09624 caenorhabdi
12	131	3.8	1422	1 YAG3_YEAST	P39712 saccharomye
13	130.5	3.8	768	1 AMV1_SAO10	P29760 saccharomye
14	130.5	3.8	1161	1 DANA_YEAST	P47179 saccharomye
15	129.5	3.8	1421	1 PCGN_HUMAN	Q14594 homo sapien
16	129.5	3.8	1467	1 AMYH_YEAST	P08540 saccharomye
17	129	3.8	1015	1 P1FX_HUMAN	Q92942 homo sapien
18	128.5	3.8	1537	1 F101_YEAST	P32768 saccharomye
19	128	3.8	767	1 AMYH_SAO10	P40465 saccharomye
20	127.5	3.7	518	1 VLZ2_HIV47	P32425 human papil
21	126.5	3.7	451	1 C7A1_HCV	P06663 turnip rin
22	126.5	3.7	1406	1 MSR2_YEAST	P42334 saccharomye
23	125.5	3.7	725	1 A1A1_YEAST	P42323 saccharomye
24	125	3.7	697	1 RYN_DROME	P95965 drosophila
25	125	3.7	1054	1 S24A_AFA1H	Q95100 arabidopsi
26	124.5	3.6	797	1 VGLX_HSV8	P28968 equine corp
27	124.5	3.6	1119	1 ALB3_CANAL	G74623 candiru alb
28	124.5	3.6	1140	1 YMR6_YEAST	Q04893 saccharomye
29	123	3.5	2096	1 HET1_MESAO	P51611 mesocricetu
30	122	3.6	472	1 SX14_DROME	P40656 drosophila
31	122	3.6	354	1 CA0N_HUMAN	Q94251 homo sapien
32	121.5	3.6	743	1 P021_HUMAN	P14859 homo sapien
33	121	3.6	676	1 EXL1_HUMAN	Q92935 homo sapien

ALIGNMENTS

RESULT 1	VS12_HEVR0	STANDARD	PERC	660 AA
AC	P29326			
DT	01-DEC-1992 (Ref. 24, Created)			
DT	01-DEC-1992 (Ref. 24, Last sequence update)			
DE	01-FEB-1994 (Ref. 28, Last annotation update)			
DE	Structural protein 2 precursor (p0p2)			
OS	Hepatitis E virus (strain Burma) (HEV)			
OC	Viruses, ssRNA positive strand viruses, no RNA stage			
OC	Hepatitis E-like viruses			
OX	NCBI TaxID 41767			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE 9202467; PubMed 190770			
RA	Tam A.W., Smith M.M., Guerra M.E., Huang C.C., Bradley D.W.,			
EA	Fry K.E., Reyes G.R.			
ET	"Hepatitis E virus (HEV): molecular cloning and sequencing of the			
RT	full-length viral genome."			
RL	Virology 185:120-131(1991).			
CC	1- FUNCTION: CONTAINS A HIGH BASIC AMINE WITH AN EXTREME SUBSTITUTED			
CC	BY EFFECTIVELY NEUTRALIZING THE RESIDUALLY CHARGED RNA			

THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement with EMBL. Send an email to EMBL@EMBL.ch.

EMBL: M73218; AAA15746.1				
DR PIR: C40778; VHWHL2				
DR InterPro: IPR024261, 372				
DR Pfam: PF03014; SP2.1				
KW Signal				
FT SIGNAL 1 19	POTENTIAL			
FT CHAIN 23 660	STRUCTURAL PROTEIN			
SQ SEQUENCE 660 AA: 76978 MW: 58420.1504 kDa				

Query Match 100.0% Score: 6408 DB ID: Length: 660
Best local similarity: 100.0% ID: 90.2% Gap: 207
Matches: 660; Conserved: 0; Mismatches: 0; Indels: 0; Gaps: 0;

Q7	1	MAKRLLELLELPKAPFETQGLGKGRGRKLSKQFWSPVQAPAFVDFHTN 60			
DB	1	MRPRLILLMLLMFLMFLAPFQSSPGRKPKSSSSFWSPVQAPAFVDFHTN 60			
Q7	61	PTAVVVAACACGPPVRCFAPTCZAWKQCAQFVAVASRPFTLGAAPLAVAFHT 120			
DB	61	PTAVVVAACACGPPVRCFAPTCZAWKQCAQFVAVASRPFTLGAAPLAVAFHT 120			
Q7	121	EVKLVKFAATLPRVYVLTGKLVKLVAVAVKSLTLLGSHHMAFAP 180			
DB	121	EVKLVKFAATLPRVYVLTGKLVKLVAVAVKSLTLLGSHHMAFAP 180			

QY 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPPHVPPT 240
 DB 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPPHVPPT 240
 QY 241 ASEVIPSEIRHYRNOGMRWSVETSGVAFFATSI VMLC IUGSLVNSYNTPTVTSALGILL 300
 DB 241 ASEVIPSEIRHYRNOGMRWSVETSGVAFFATSI VMLC IUGSLVNSYNTPTVTSALGILL 300
 QY 301 DFALFELFRNLTPGNTNTRVRSYSTARHLRRCAGDTAELTTTATIRMKOLYFTSTNG 360
 DB 301 DFALFELFRNLTPGNTNTRVRSYSTARHLRRCAGDTAELTTTATIRMKOLYFTSTNG 360
 QY 361 VGEIGRGIAITLFLNLAITLGGTPTFISSAGGZLYSKVSVNSANSEPIVKLVISVENAQ 420
 DB 361 VGEIGRGIAITLFLNLAITLGGTPTFISSAGGZLYSKVSVNSANSEPIVKLVISVENAQ 420
 QY 421 QPKGIATPHDIDGFSRVVIGQYDNQHEQDPPTSPAPSPFPFVLFPANDVLMLSLTAABY 480
 DB 421 QPKGIATPHDIDGFSRVVIGQYDNQHEQDPPTSPAPSPFPFVLFPANDVLMLSLTAABY 480
 QY 481 DQSTYSSSTGVVYSVSVTLNVNATGACAVAPSLDWTKVTLDGRPTSTTOOYSKTFFVLP 540
 DB 481 DQSTYSSSTGVVYSVSVTLNVNATGACAVAPSLDWTKVTLDGRPTSTTOOYSKTFFVLP 540
 QY 541 LPCKISFWFAGTTRKAGYPYNTTASQGLIVFNAAHFPVATSTYVTSLSGAGPVVISAVAV 600
 DB 541 LPCKISFWFAGTTRKAGYPYNTTASQGLIVFNAAHFPVATSTYVTSLSGAGPVVISAVAV 600
 QY 601 LAPHSALELFDITLQYPARAHITDPTDPCPPFPGIQTAPFSTVAHLQPKMKPVKPRFRI 660
 DB 601 LAPHSALELFDITLQYPARAHITDPTDPCPPFPGIQTAPFSTVAHLQPKMKPVKPRFRI 660

RESULT 2

VST2_HEVPA STANDARD: PPT: 660 AA
 AC P33426;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Structural protein 2 precursor (ORF2).
 OS Hepatitis E virus (strain Pakistan) (HEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses
 OX NCBI_TaxID=33774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92115700; PubMed=1731327;
 RA Tsarev S.A., Emerson S.O., Reyes G.R., Tsareva T.S., Lecters L.J.,
 RA Malik I.A., Iqbal M., Purcell R.H.;
 R1 "Characterization of a prototype strain of hepatitis E virus."
 RL Proc Natl Acad Sci U S A 89:550-563(1992)
 CC 1-1 FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
 CC THAT IT MAY BE INVOLVED IN THE ENCAPSULATION OF THE GENOMIC RNA
 CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED PNA
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC
 DR EMBL: M80584; AAA45727.1;
 DR InterPro: IPI004261; SP2.
 DR Pfam: PF03014; SP2; 1.
 KW Signal.
 FT SIGNAL. 1 22 BY SIMILARITY.
 FT CHAIN. 23 660 STRUCTURAL PROTEIN 2.
 FT SPOTEN-F 660 AA; 70080 MW; 80856-530-FM45F33-CP064;

Query Match 99.5%; Score 3391; DB 1; Length 660;
 Best Local Similarity 99.4%; Prob. No. 2.2e-206;
 Matches 656; Conservative 1; Mismatches 5; Indels 6; Gaps 0
 QY 1 MEPPPIIIIIIMELPMI PAPPQQPSPPRRRRSGSGSGGSEWDRVDSQFFPAIYIHPIN 60
 DB 1 MEPPPIIIIIIMELPMI PAPPQQPSPPRRRRSGSGSGGSEWDRVDSQFFPAIYIHPIN 60
 QY 61 PFADVTAAACAGCPVPQPAHPHLOGSAWPGCAQHPAVASHPRPPIIAAAILAVAIADHTP 120
 DB 61 PFADVTAAACAGCPVPQPAHPHLOGSAWPGCAQHPAAASHPRPPIIAAAILAVAIADHTP 120
 QY 121 PVQVDPSPGATIPQYNI STSPITSSVATGNI VI VAAPI SDI PTQNTNTHIMAFAS 180
 DB 121 PVQVDPSPGATIPQYNI STSPITSSVATGNI VI VAAPI SDI PTQNTNTHIMAFAS 180
 QY 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPPHVPPT 240
 DB 181 NYAGYVPVAPATIPYPPPIVNVAVGGYALISSEWPGTITITPTISVFMNSITSTVPPHVPPT 240
 QY 241 ASEVIPSEIRHYRNOGMRWSVETSGVAFFATSI VMLC IUGSLVNSYNTPTVTSALGILL 300
 DB 241 ASEVIPSEIRHYRNOGMRWSVETSGVAFFATSI VMLC IUGSLVNSYNTPTVTSALGILL 300
 QY 301 DFALFELFRNLTPGNTNTRVRSYSTARHLRRCAGDTAELTTTATIRMKOLYFTSTNG 360
 DB 301 DFALFELFRNLTPGNTNTRVRSYSTARHLRRCAGDTAELTTTATIRMKOLYFTSTNG 360
 QY 361 VGEIGRGIAITLFLNLAITLGGTPTFISSAGGZLYSKVSVNSANSEPIVKLVISVENAQ 420
 DB 361 VGEIGRGIAITLFLNLAITLGGTPTFISSAGGZLYSKVSVNSANSEPIVKLVISVENAQ 420
 QY 421 QPKGIATPHDIDGFSRVVIGQYDNQHEQDPPTSPAPSPFPFVLFPANDVLMLSLTAABY 480
 DB 421 QPKGIATPHDIDGFSRVVIGQYDNQHEQDPPTSPAPSPFPFVLFPANDVLMLSLTAABY 480
 QY 481 DQSTYSSSTGVVYSVSVTLNVNATGACAVAPSLDWTKVTLDGRPTSTTOOYSKTFFVLP 540
 DB 481 DQSTYSSSTGVVYSVSVTLNVNATGACAVAPSLDWTKVTLDGRPTSTTOOYSKTFFVLP 540
 QY 541 LPCKISFWFAGTTRKAGYPYNTTASQGLIVFNAAHFPVATSTYVTSLSGAGPVVISAVAV 600
 DB 541 LPCKISFWFAGTTRKAGYPYNTTASQGLIVFNAAHFPVATSTYVTSLSGAGPVVISAVAV 600
 QY 601 LAPHSALELFDITLQYPARAHITDPTDPCPPFPGIQTAPFSTVAHLQPKMKPVKPRFRI 660
 DB 601 LAPHSALELFDITLQYPARAHITDPTDPCPPFPGIQTAPFSTVAHLQPKMKPVKPRFRI 660

RESULT 3

VST2_HEVMY STANDARD: PPT: 660 AA.
 AC Q04611;
 DT 01-OCT-1994 (Rel. 27, Created)
 DT 01-OCT-1994 (Rel. 27, Last sequence update)
 DT 01-NOV-1995 (Rel. 42, Last annotation update)
 DE Structural protein 2 precursor (ORF2).
 OS Hepatitis E virus (strain Myanmar) (HEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=41769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93227577; PubMed 9470471;
 RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
 RA Ekihiro T., Waki K.;
 R1 "Sequence and gene structure of the hepatitis E virus isolated from
 R1 Myanmar."
 RL Virus Genes 7:95-109(1994).
 CC 1-1 FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
 CC THAT IT MAY BE INVOLVED IN THE ENCAPSULATION OF THE GENOMIC RNA
 CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED PNA
 CC

CC This SWISS prot entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre>
 CC or send an email to license@ebi.ac.uk).

DB EMBL: D10450; GenBank: U001174.1;
 DB UniProt: P000426.1; SP2;
 DB Pfam: PF04014; SP2; 1;
 KW SIGNAL;
 FT CHAIN; 1..22 BY SIMILARITY;
 FT CHAIN; 23..660 STRUCTURAL PROTEIN 2;
 SQ SEQUENCE 660 AA; 70998 MW; 3482044EA155C6552 CRC64;

Query Match 98.8%; Score 3367; DB 1; Length 660;
 Best Local Similarity 98.9%; Pred No 1e-204;
 Matches 653; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRPRILLMLPMLPAPPGGSGRRGRSGSGGFWGRPVISQFFAIPYTHPTN 60
 DB 1 MRPRILLMLPMLPAPPGGSGRRGRSGSGGFWGRPVISQFFAIPYTHPTN 60
 QY 61 PFAPVTAAGAGPRVGPAPRPGSGAMRQGRVAVASRRRPTTAGAAPLTAVAPADTP 120
 DB 61 PFAPVTAAGAGPRVGPAPRPGSGAMRQGRVAVASRRRPTTAGAAPLTAVAPADTP 120
 QY 121 PVPVUSGAILRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 180
 DB 121 PVPVUSGAILRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 180
 QY 181 NYAQYRVKATIRYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 240
 DB 181 NYAQYRVKATIRYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 240
 QY 241 ASELYPSEERLYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 300
 DB 241 ASELYPSEERLYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 300
 QY 301 DPALEERPLNLTPTNTNVSYSSTAPRRPGAGCTAFITTAATRFMKDLYFTSTNG 360
 DB 301 DPALEERPLNLTPTNTNVSYSSTAPRRPGAGCTAFITTAATRFMKDLYFTSTNG 360
 QY 361 VVEIQRGIALFLNIADTLGGLPTELISAGGGLYSPRVVSANGRPYTKLYTSVENAQ 420
 DB 361 VVEIQRGIALFLNIADTLGGLPTELISAGGGLYSPRVVSANGRPYTKLYTSVENAQ 420
 QY 421 QIKGIALPHDILGDSRWVIGQYQNGRQGRPTSPAPSRPFVLRANGVLSLTAAY 480
 DB 421 QIKGIALPHDILGDSRWVIGQYQNGRQGRPTSPAPSRPFVLRANGVLSLTAAY 480
 QY 481 QSTVGSSTAPVYVSSTLVNVAIGAVARSIDWTKVTLGRPLSTIQQYKSTFEVLP 540
 DB 481 QSTVGSSTAPVYVSSTLVNVAIGAVARSIDWTKVTLGRPLSTIQQYKSTFEVLP 540
 QY 541 LRKGLSEWAGIKRAGHYNTNTASDQLLVENAGRRVALSYITSLAGPVSISAVAV 600
 DB 541 LRKGLSEWAGIKRAGHYNTNTASDQLLVENAGRRVALSYITSLAGPVSISAVAV 600
 QY 601 LAPHSALALETLNYPAPATFDDPFPYPPRPLAQCAFQSTVAEIQRLKMKVGKTRRL 660
 DB 601 LAPHSALALETLNYPAPATFDDPFPYPPRPLAQCAFQSTVAEIQRLKMKVGKTRRL 660

RESULT 1
 VS12 REVME
 ID VS12_REVME STANDARD; PRI: 65% AA
 AC 60560
 DI OCT 1993 (601, 27, Created)
 DI OCT 1993 (601, 27, Last sequence update)
 DI FEB 1994 (601, 28, Last annotation update)
 DE Structural protein 2 precursor.

CC Hepatitis E virus (strain R410a) (HVE)
 CC Virus; ssRNA positive strand viruses; no DNA stage;
 CC Hepatitis E-like viruses;
 CC NCBI TaxID 41766;
 RN 111
 RI 1
 RI 2
 RI 3
 RI 4
 RI 5
 RI 6
 RI 7
 RI 8
 RI 9
 RI 10
 RI 11
 RI 12
 RI 13
 RI 14
 RI 15
 RI 16
 RI 17
 RI 18
 RI 19
 RI 20
 RI 21
 RI 22
 RI 23
 RI 24
 RI 25
 RI 26
 RI 27
 RI 28
 RI 29
 RI 30
 RI 31
 RI 32
 RI 33
 RI 34
 RI 35
 RI 36
 RI 37
 RI 38
 RI 39
 RI 40
 RI 41
 RI 42
 RI 43
 RI 44
 RI 45
 RI 46
 RI 47
 RI 48
 RI 49
 RI 50
 RI 51
 RI 52
 RI 53
 RI 54
 RI 55
 RI 56
 RI 57
 RI 58
 RI 59
 RI 60
 RI 61
 RI 62
 RI 63
 RI 64
 RI 65
 RI 66
 RI 67
 RI 68
 RI 69
 RI 70
 RI 71
 RI 72
 RI 73
 RI 74
 RI 75
 RI 76
 RI 77
 RI 78
 RI 79
 RI 80
 RI 81
 RI 82
 RI 83
 RI 84
 RI 85
 RI 86
 RI 87
 RI 88
 RI 89
 RI 90
 RI 91
 RI 92
 RI 93
 RI 94
 RI 95
 RI 96
 RI 97
 RI 98
 RI 99
 RI 100
 RI 101
 RI 102
 RI 103
 RI 104
 RI 105
 RI 106
 RI 107
 RI 108
 RI 109
 RI 110
 RI 111
 RI 112
 RI 113
 RI 114
 RI 115
 RI 116
 RI 117
 RI 118
 RI 119
 RI 120
 RI 121
 RI 122
 RI 123
 RI 124
 RI 125
 RI 126
 RI 127
 RI 128
 RI 129
 RI 130
 RI 131
 RI 132
 RI 133
 RI 134
 RI 135
 RI 136
 RI 137
 RI 138
 RI 139
 RI 140
 RI 141
 RI 142
 RI 143
 RI 144
 RI 145
 RI 146
 RI 147
 RI 148
 RI 149
 RI 150
 RI 151
 RI 152
 RI 153
 RI 154
 RI 155
 RI 156
 RI 157
 RI 158
 RI 159
 RI 160
 RI 161
 RI 162
 RI 163
 RI 164
 RI 165
 RI 166
 RI 167
 RI 168
 RI 169
 RI 170
 RI 171
 RI 172
 RI 173
 RI 174
 RI 175
 RI 176
 RI 177
 RI 178
 RI 179
 RI 180
 RI 181
 RI 182
 RI 183
 RI 184
 RI 185
 RI 186
 RI 187
 RI 188
 RI 189
 RI 190
 RI 191
 RI 192
 RI 193
 RI 194
 RI 195
 RI 196
 RI 197
 RI 198
 RI 199
 RI 200
 RI 201
 RI 202
 RI 203
 RI 204
 RI 205
 RI 206
 RI 207
 RI 208
 RI 209
 RI 210
 RI 211
 RI 212
 RI 213
 RI 214
 RI 215
 RI 216
 RI 217
 RI 218
 RI 219
 RI 220
 RI 221
 RI 222
 RI 223
 RI 224
 RI 225
 RI 226
 RI 227
 RI 228
 RI 229
 RI 230
 RI 231
 RI 232
 RI 233
 RI 234
 RI 235
 RI 236
 RI 237
 RI 238
 RI 239
 RI 240
 RI 241
 RI 242
 RI 243
 RI 244
 RI 245
 RI 246
 RI 247
 RI 248
 RI 249
 RI 250
 RI 251
 RI 252
 RI 253
 RI 254
 RI 255
 RI 256
 RI 257
 RI 258
 RI 259
 RI 260
 RI 261
 RI 262
 RI 263
 RI 264
 RI 265
 RI 266
 RI 267
 RI 268
 RI 269
 RI 270
 RI 271
 RI 272
 RI 273
 RI 274
 RI 275
 RI 276
 RI 277
 RI 278
 RI 279
 RI 280
 RI 281
 RI 282
 RI 283
 RI 284
 RI 285
 RI 286
 RI 287
 RI 288
 RI 289
 RI 290
 RI 291
 RI 292
 RI 293
 RI 294
 RI 295
 RI 296
 RI 297
 RI 298
 RI 299
 RI 300
 RI 301
 RI 302
 RI 303
 RI 304
 RI 305
 RI 306
 RI 307
 RI 308
 RI 309
 RI 310
 RI 311
 RI 312
 RI 313
 RI 314
 RI 315
 RI 316
 RI 317
 RI 318
 RI 319
 RI 320
 RI 321
 RI 322
 RI 323
 RI 324
 RI 325
 RI 326
 RI 327
 RI 328
 RI 329
 RI 330
 RI 331
 RI 332
 RI 333
 RI 334
 RI 335
 RI 336
 RI 337
 RI 338
 RI 339
 RI 340
 RI 341
 RI 342
 RI 343
 RI 344
 RI 345
 RI 346
 RI 347
 RI 348
 RI 349
 RI 350
 RI 351
 RI 352
 RI 353
 RI 354
 RI 355
 RI 356
 RI 357
 RI 358
 RI 359
 RI 360
 RI 361
 RI 362
 RI 363
 RI 364
 RI 365
 RI 366
 RI 367
 RI 368
 RI 369
 RI 370
 RI 371
 RI 372
 RI 373
 RI 374
 RI 375
 RI 376
 RI 377
 RI 378
 RI 379
 RI 380
 RI 381
 RI 382
 RI 383
 RI 384
 RI 385
 RI 386
 RI 387
 RI 388
 RI 389
 RI 390
 RI 391
 RI 392
 RI 393
 RI 394
 RI 395
 RI 396
 RI 397
 RI 398
 RI 399
 RI 400
 RI 401
 RI 402
 RI 403
 RI 404
 RI 405
 RI 406
 RI 407
 RI 408
 RI 409
 RI 410
 RI 411
 RI 412
 RI 413
 RI 414
 RI 415
 RI 416
 RI 417
 RI 418
 RI 419
 RI 420
 RI 421
 RI 422
 RI 423
 RI 424
 RI 425
 RI 426
 RI 427
 RI 428
 RI 429
 RI 430
 RI 431
 RI 432
 RI 433
 RI 434
 RI 435
 RI 436
 RI 437
 RI 438
 RI 439
 RI 440
 RI 441
 RI 442
 RI 443
 RI 444
 RI 445
 RI 446
 RI 447
 RI 448
 RI 449
 RI 450
 RI 451
 RI 452
 RI 453
 RI 454
 RI 455
 RI 456
 RI 457
 RI 458
 RI 459
 RI 460
 RI 461
 RI 462
 RI 463
 RI 464
 RI 465
 RI 466
 RI 467
 RI 468
 RI 469
 RI 470
 RI 471
 RI 472
 RI 473
 RI 474
 RI 475
 RI 476
 RI 477
 RI 478
 RI 479
 RI 480
 RI 481
 RI 482
 RI 483
 RI 484
 RI 485
 RI 486
 RI 487
 RI 488
 RI 489
 RI 490
 RI 491
 RI 492
 RI 493
 RI 494
 RI 495
 RI 496
 RI 497
 RI 498
 RI 499
 RI 500
 RI 501
 RI 502
 RI 503
 RI 504
 RI 505
 RI 506
 RI 507
 RI 508
 RI 509
 RI 510
 RI 511
 RI 512
 RI 513
 RI 514
 RI 515
 RI 516
 RI 517
 RI 518
 RI 519
 RI 520
 RI 521
 RI 522
 RI 523
 RI 524
 RI 525
 RI 526
 RI 527
 RI 528
 RI 529
 RI 530
 RI 531
 RI 532
 RI 533
 RI 534
 RI 535
 RI 536
 RI 537
 RI 538
 RI 539
 RI 540
 RI 541
 RI 542
 RI 543
 RI 544
 RI 545
 RI 546
 RI 547
 RI 548
 RI 549
 RI 550
 RI 551
 RI 552
 RI 553
 RI 554
 RI 555
 RI 556
 RI 557
 RI 558
 RI 559
 RI 560
 RI 561
 RI 562
 RI 563
 RI 564
 RI 565
 RI 566
 RI 567
 RI 568
 RI 569
 RI 570
 RI 571
 RI 572
 RI 573
 RI 574
 RI 575
 RI 576
 RI 577
 RI 578
 RI 579
 RI 580
 RI 581
 RI 582
 RI 583
 RI 584
 RI 585
 RI 586
 RI 587
 RI 588
 RI 589
 RI 590
 RI 591
 RI 592
 RI 593
 RI 594
 RI 595
 RI 596
 RI 597
 RI 598
 RI 599
 RI 600
 RI 601
 RI 602
 RI 603
 RI 604
 RI 605
 RI 606
 RI 607
 RI 608
 RI 609
 RI 610
 RI 611
 RI 612
 RI 613
 RI 614
 RI 615
 RI 616
 RI 617
 RI 618
 RI 619
 RI 620
 RI 621
 RI 622
 RI 623
 RI 624
 RI 625
 RI 626
 RI 627
 RI 628
 RI 629
 RI 630
 RI 631
 RI 632
 RI 633
 RI 634
 RI 635
 RI 636
 RI 637
 RI 638
 RI 639
 RI 640
 RI 641
 RI 642
 RI 643
 RI 644
 RI 645
 RI 646
 RI 647
 RI 648
 RI 649
 RI 650
 RI 651
 RI 652
 RI 653
 RI 654
 RI 655
 RI 656
 RI 657
 RI 658
 RI 659
 RI 660
 RI 661
 RI 662
 RI 663
 RI 664
 RI 665
 RI 666
 RI 667
 RI 668
 RI 669
 RI 670
 RI 671
 RI 672
 RI 673
 RI 674
 RI 675
 RI 676
 RI 677
 RI 678
 RI 679
 RI 680
 RI 681
 RI 682
 RI 683
 RI 684
 RI 685
 RI 686
 RI 687
 RI 688
 RI 689
 RI 690
 RI 691
 RI 692
 RI 693
 RI 694
 RI 695
 RI 696
 RI 697
 RI 698
 RI 699
 RI 700
 RI 701
 RI 702
 RI 703
 RI 704
 RI 705
 RI 706
 RI 707
 RI 708
 RI 709
 RI 710
 RI 711
 RI 712
 RI 713
 RI 714
 RI 715
 RI 716
 RI 717
 RI 718
 RI 719
 RI 720
 RI 721
 RI 722
 RI 723
 RI 724
 RI 725
 RI 726
 RI 727
 RI 728
 RI 729
 RI 730
 RI 731
 RI 732
 RI 733
 RI 734
 RI 735
 RI 736
 RI 737
 RI 738
 RI 739
 RI 740
 RI 741
 RI 742
 RI 743
 RI 744
 RI 745
 RI 746
 RI 747
 RI 748
 RI 749
 RI 750
 RI 751
 RI 752
 RI 753
 RI 754
 RI 755
 RI 756
 RI 757
 RI 758
 RI 759
 RI 760
 RI 761
 RI 762
 RI 763
 RI 764
 RI 765
 RI 766
 RI 767
 RI 768
 RI 769
 RI 770
 RI 771
 RI 772
 RI 773
 RI 774
 RI 775
 RI 776
 RI 777
 RI 778
 RI 779
 RI 780
 RI 781
 RI 782
 RI 783
 RI 784
 RI 785
 RI 786
 RI 787
 RI 788
 RI 789
 RI 790
 RI 791
 RI 792
 RI 793
 RI 794
 RI 795
 RI 796
 RI 797
 RI 798
 RI 799
 RI 800
 RI 801
 RI 802
 RI 803
 RI 804
 RI 805
 RI 806
 RI 807
 RI 808
 RI 809
 RI 810
 RI 811
 RI 812
 RI 813
 RI 814
 RI 815
 RI 816
 RI 817
 RI 818
 RI 819
 RI 820
 RI 821
 RI 822
 RI 823
 RI 824
 RI 825
 RI 826
 RI 827
 RI 828
 RI 829
 RI 830
 RI 831
 RI 832
 RI 833
 RI 834
 RI 835
 RI 836
 RI 837
 RI 838
 RI 839
 RI 840
 RI 841
 RI 842
 RI 843
 RI 844
 RI 845
 RI 846
 RI 847
 RI 848
 RI 849
 RI 850
 RI 851
 RI 852
 RI 853
 RI 854
 RI 855
 RI 856
 RI 857
 RI 858
 RI 859
 RI 860
 RI 861
 RI 862
 RI 863
 RI 864
 RI 865
 RI 866
 RI 867
 RI 868
 RI 869
 RI 870
 RI 871
 RI 872
 RI 873
 RI 874
 RI 875
 RI 876
 RI 877
 RI 878
 RI 879
 RI 880
 RI 881
 RI 882
 RI 883
 RI 884
 RI 885
 RI 886
 RI 887
 RI 888
 RI 889
 RI 890
 RI 891
 RI 892
 RI 893
 RI 894
 RI 895
 RI 896
 RI 897
 RI 898
 RI 899
 RI 900
 RI 901
 RI 902
 RI 903
 RI 904
 RI 905
 RI 906
 RI 907
 RI 908
 RI 909
 RI 910
 RI 911
 RI 912
 RI 913
 RI 914
 RI 915
 RI 916
 RI 917
 RI 918
 RI 919
 RI 920
 RI 921
 RI 922
 RI 923
 RI 924
 RI 925
 RI 926
 RI 927
 RI 928
 RI 929
 RI 930
 RI 931
 RI 932
 RI 933
 RI 934
 RI 935
 RI 936
 RI 937
 RI 938
 RI 939
 RI 940
 RI 941
 RI 942
 RI 943
 RI 944
 RI 945
 RI 946
 RI 947
 RI 948
 RI 949
 RI 950
 RI 951
 RI 952
 RI 953
 RI 954
 RI 955
 RI 956
 RI 957
 RI 958
 RI 959
 RI 960
 RI 961
 RI 962
 RI 963
 RI 964
 RI 965
 RI 966
 RI 967
 RI 968
 RI 969
 RI 970
 RI 971
 RI 972
 RI 973
 RI 974
 RI 975
 RI 976
 RI 977
 RI 978
 RI 979
 RI 980
 RI 981
 RI 982
 RI 983
 RI 984
 RI 985
 RI 986
 RI 987
 RI 988
 RI 989
 RI 990
 RI 991
 RI 992
 RI 993
 RI 994
 RI 995
 RI 996
 RI 997
 RI 998
 RI 999
 RI 1000

CC This SWISS prot entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre>
 CC or send an email to license@ebi.ac.uk).

DB EMBL: M74506; GenBank: U001174.1;
 DB UniProt: P000426.1; SP2;
 DB Pfam: PF04014; SP2; 1;
 KW SIGNAL;
 FT CHAIN; 1..22 BY SIMILARITY;
 FT CHAIN; 23..660 STRUCTURAL PROTEIN 2;
 SQ SEQUENCE 660 AA; 70998 MW; 3482044EA155C6552 CRC64;

Query Match 98.8%; Score 3367; DB 1; Length 660;
 Best Local Similarity 98.9%; Pred No 1e-204;
 Matches 653; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRPRILLMLPMLPAPPGGSGRRGRSGSGGFWGRPVISQFFAIPYTHPTN 60
 DB 1 MRPRILLMLPMLPAPPGGSGRRGRSGSGGFWGRPVISQFFAIPYTHPTN 60
 QY 61 PFAPVTAAGAGPRVGPAPRPGSGAMRQGRVAVASRRRPTTAGAAPLTAVAPADTP 120
 DB 61 PFAPVTAAGAGPRVGPAPRPGSGAMRQGRVAVASRRRPTTAGAAPLTAVAPADTP 120
 QY 121 PVPVUSGAILRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 180
 DB 121 PVPVUSGAILRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 180
 QY 181 NYAQYRVKATIRYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 240
 DB 181 NYAQYRVKATIRYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 240
 QY 241 ASELYPSEERLYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 300
 DB 241 ASELYPSEERLYRQYNLSISPLSSVATGTLNLYAAPLSPLPPLQKINTHIMAEAS 300
 QY 301 DPALEERPLNLTPTNTNVSYSSTAPRRPGAGCTAFITTAATRFMKDLYFTSTNG 360
 DB 301 DPALEERPLNLTPTNTNVSYSSTAPRRPGAGCTAFITTAATRFMKDLYFTSTNG 360
 QY 361 VVEIQRGIALFLNIADTLGGLPTELISAGGGLYSPRVVSANGRPYTKLYTSVENAQ 420
 DB 361 VVEIQRGIALFLNIADTLGGLPTELISAGGGLYSPRVVSANGRPYTKLYTSVENAQ 420
 QY 421 QIKGIALPHDILGDSRWVIGQYQNGRQGRPTSPAPSRPFVLRANGVLSLTAAY 480
 DB 421 QIKGIALPHDILGDSRWVIGQYQNGRQGRPTSPAPSRPFVLRANGVLSLTAAY 480
 QY 481 QSTVGSSTAPVYVSSTLVNVAIGAVARSIDWTKVTLGRPLSTIQQYKSTFEVLP 540
 DB 481 QSTVGSSTAPVYVSSTLVNVAIGAVARSIDWTKVTLGRPLSTIQQYKSTFEVLP 540
 QY 541 LRKGLSEWAGIKRAGHYNTNTASDQLLVENAGRRVALSYITSLAGPVSISAVAV 600
 DB 541 LRKGLSEWAGIKRAGHYNTNTASDQLLVENAGRRVALSYITSLAGPVSISAVAV 600
 QY 601 LAPHSALALETLNYPAPATFDDPFPYPPRPLAQCAFQSTVAEIQRLKMKVGKTRRL 660
 DB 601 LAPHSALALETLNYPAPATFDDPFPYPPRPLAQCAFQSTVAEIQRLKMKVGKTRRL 660

RESULT 1
 VS12 REVME
 ID VS12_REVME STANDARD; PRI: 65% AA
 AC 60560
 DI OCT 1993 (601, 27,

```

QY 541 LRGKLSFWFACTHAGHPYNNYNIASQILVFNAAHHPVAISYHLSIAGHPVSIASAVAV 599
DB 540 LRGKLSWEAGTTKAGYPYNNYNTASQILVFNAAHHPVAISYHLSIAGHPVSIASAVAV 599
QY 601 LAPHSAALLLEDLDYPARAHITDDFCPCPCRPGLGLOFCAPQSTVAFLOPKMKVGGKTRFL 660
DB 600 LAPPASALLLEDYFHYNGAPAHITDDFCPCPCRPGLGLOFCAPQSTVAFLOPKMKVGGKTRFL 659

RESULT 5
ID VST2 HEVRH STANDARD; PRT; 485 AA.
AC Q00270;
DT 01-OCT-1994 (Rel. 27, Created)
DI 01-OCT-1994 (Rel. 27, Last sequence update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Hecus) (HEV)
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
CC Hepatitis E-like viruses.
GX NCBI_taxid=31766;
FN [1]
RP SEQUENCE FROM N A
RX MEDLINE=92261377; PubMed=1484074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Ogasawara S.,
RA Wang C.-K., Shikata T., Ichikawa M., Iikihisa T., Mizuno K.,
RA Wino K.M.;
RT "Hepatitis E virus: cDNA cloning and expression."
RL Microbiol Immunol 36:67-79(1992)
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@sib-sib.ch)
CC
DP EMBL: p00274; RAA209010.1; -
DR InterPro: IPR004261; SP2.
DR Pfam: PF03014; SP2; 1.
FT NON_TER 1
FT NON_TER 485
SU SEQUENCE 485 AA: 52417 MW: 54592727.2747494 CACG64;

Query Match
Best Local Similarity: 99.0% Prod No: 64147;
Matches 480, Conservative 1, Mismatches 4, Indels 6, Gaps 0;

QY 133 KROYNLSTSPITSSVATGNIIVLYAAPI SPLPLDQGTNTHIMATPASNYAYVVPAT1 192
DB 1 KROYNLSTSPITSSVATGNIIVLYAAPLSPLPLDQGTNTHIMATEASNYAYVVRAT1 60
QY 193 KYPVLVNAVGGVAISSEWPGTTITFISVNMNSIISILVPLVGVGTASELVLISEK1H 252
DB 61 KYPVLVNAVGGVAISSEWPGTTITFISVNMNSIISILVPLVGVGTASELVLISEK1H 129
QY 253 YPNQWPSVETSQVAFEFATSGIVMLCIGSLVNSVNTPTGTGALLDFALEFPNLT 312
DB 121 YPNQWPSVETSQVAFEFATSGIVMLCIGSLVNSVNTPTGTGALLDFALEFPNLT 180
QY 313 PNTINTVSPYSSTAPHPPIPGAGCTAPITTTAATPFPMALVETSTNGVCEIGCGIALTL 372
DB 181 PNTINTVSPYSSTAPHPPIPGAGCTAPITTTAATPFPMALVETSTNGVCEIGCGIALTL 240
QY 373 PNLACTLIGGIPETLLESAAGGLFYGFVFWVSARHETVTVTSVTHAAGCFGATHTH 430
DB 241 PNLACTLIGGIPETLLESAAGGLFYGFVFWVSARHETVTVTSVTHAAGCFGATHTH 300
QY 433 LKFSVVIQYDYNHCFQHPPTSPAPSPFVLPANFVLMVLSLTAEEVQSTYSGSTGV 492
DB 301 LKFSVVIQYDYNHCFQHPPTSPAPSPFVLPANFVLMVLSLTAEEVQSTYSGSTGV 360

```

```

QY 493 VVSQSVLIVVAIAGAVAPASITWTVTITGDPISITIGVSPSTFVILHFKWISWEAFT 552
DB 361 VVSQSVLIVVAIAGAVAPASITWTVTITGDPISITIGVSPSTFVILHFKWISWEAFT 420
QY 553 TRAGYPYNYNTASDOLLVFNAAHHPVAISYHLSIAGHPVSIASAVAVIAPHSAIALLD 612
DB 421 TRAGYPYNYNTASDOLLVFNAAHHPVAISYHLSIAGHPVSIASAVAVIAPHSAIALLD 480
QY 613 TLDYP 617
DB 481 TLDYP 485

RESULT 6
ID EGT2 YEAST STANDARD; PRT; 1041 AA.
AC P42835;
DT 01-NOV-1995 (Rel. 32, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DE EGT2 protein precursor (Early G1 transcript 2).
OS Saccharomyces cerevisiae (Baker's yeast).
CC Fink G.R., Fink G.R., Fink G.R., Fink G.R., Fink G.R., Fink G.R.,
CC Saccharomyces cerevisiae (Baker's yeast).
GX NCBI_taxid=4942;
FN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=52886 / FY1676;
RX MEDLINE=95373280; PubMed 7645447;
RA Matsumoto M., Niwa M., Iwase T., M. Gullerud C.J.
RT "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV
RT identifies the EGT2, YAK8 and EPE1 loci. Five new open reading
RT frames."
RL Yeast 11:567-572(1995).
FN [2]
RP SEQUENCE FROM N.A. AND FUNCTION.
RX MEDLINE=96251274; PubMed 8668141;
RA Krasnowski P., Krasnowski P., Krasnowski P., Krasnowski P.,
RT "EGT2 gene transcription is induced predominantly by Swi5 in early
RT G1."
RL Mol. Cell. Biol. 16:3264-3274(1996).
CC -1- FUNCTION: SEEMS TO BE INVOLVED IN THE PREPARED TIMING OF CELL
CC SEPARATION AFTER CYTOKINESIS. AN INHIBITION OF MITOSIS DAUGHTER
CC CELLS IS DELAYED. COULD EITHER BE AN INTRINSIC NECESSARY FOR G1/ANAS-
CC DEGRADATION OF THE CELL WALL AT THE PROXIMAL END OF THE METABOLIC
CC DAUGHTER CELLS. A RESEMBLANCE TO THE PROXIMAL END OF THE METABOLIC
CC STEP. EXCLUSIVELY EXPRESSED BETWEEN THE END OF MITOSIS AND EARLY
CC G1. INACTIVATED BEFORE CELLS PASS START.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@sib-sib.ch)
CC
EMBL: 246259; CAA86371.1; -
DB EMBL: 271602; CAA96259.1; -
DB PIR: S55862; S55862.
DB SCV: S0005271; EGT2.
DR GO: G0000277; Cell wall (source: Fungi); IDA.
DR GO: G0000277; Cell wall (source: Fungi); IMP.
KW Glycoprotein; Receptor; Signal; Cell cycle; Cell division; Mitosis.
FT CHAIN 21 1041 EGT2 PROTEIN.
FT DOMAIN 200 203 POLY-SER.
FT DOMAIN 381 384 POLY-SER.
FT DOMAIN 493 496 POLY-SER.
FT DOMAIN 493 496 POLY-SER.
FT DOMAIN 586 589 POLY-THR.
FT REPEAT 457 492 1-1.

```


QY 475 TTAAPYQSYVSSSTGHWVVSUSVIEVNVALEALAVAPSLDWEVE-----LLE 523
 DB 1213 -TATVSSSTGPTTASNPSTITGSIENAIFFPAVAVSEPINHTLITIGTAALNYVFLVQ 1271
 QY 524 RPLSTIQOYSKIFFVLPLRGKLSPEAGITTKAGYPNYNTASTQLIVENAGHRVAIST 583
 DB 1272 NPLSSAQQ--IFNPLPLVLK-----YFNSNTSSE--LDNSICE--LST 1307
 QY 584 YTTSLCAG-----PVSISAVAVL 601
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332

RESULT 11

YS89_CABEL
 ID YS89_CABEL STANDARD: PRT: 3178 AA.
 AC Q06434.206434.206434
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 24-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein ZK945.9 in chromosome 11
 GN ZK945.9/ZK945.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ehabditida; Ehabditidae;
 OC Rhabditidae; Peloderidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Wilkinson-Sproat J.;
 RA Submitter (FEB 1995); to: Dr. FMEL, GenBank, NBR: 017805
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 BL Submitted (DEC 2001) to the EMBL/GenBank/CCRF databases
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -2- SIMILARITY: Contains 1 GFS domain.
 CC -3- SIMILARITY: Contains 1 PLAT domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (<http://www.ebi.ac.uk/submit/submit.html>)
 CC or send an email to license@ebi.scrib.ch.
 CC
 DR FMRL: Z48544; CAB70192.1; YCINFP
 DR FMRL: Z48542; CAB70192.1; YCINFP
 DR EMBL: Z48542; CAB70201.1; YCINFP
 DR WormPep: ZK945.9; CE25697.
 DR InterPro: IPR002111; Cat_channel_TripL.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001024; Lipoxigenase_LH2.
 DR InterPro: IPR003915; PKD_2
 DR InterPro: IPR000203; PKD_cys_rich.
 DR Pfam: PF01825; GFS_1.
 DR Pfam: PF00520; Ion_trans_1.
 DR Pfam: PF01477; PLAT_1.
 DR PRINTS: PR01433; POLYCYSTIN2.
 DR SMART: SM00303; GPS_1.
 DR SMART: SM00304; LH2_1.
 DR PROSITE: PS00095; PLAT; 1.
 KW Hypothetical protein, Transmembrane.
 FT TRANSMEM 13 30 POTENTIAL
 FT TRANSMEM 51 73 POTENTIAL
 FT TRANSMEM 2129 2161 POTENTIAL
 FT TRANSMEM 2348 2367 POTENTIAL
 FT TRANSMEM 2390 2412 POTENTIAL
 FT TRANSMEM 2451 2468 POTENTIAL
 FT TRANSMEM 2483 2505 POTENTIAL
 FT TRANSMEM 2567 2589 POTENTIAL

FT TRANSMEM 2946 2958 POTENTIAL
 FT TRANSMEM 2961 2961 POTENTIAL
 FT TRANSMEM 2976 2998 POTENTIAL
 FT TRANSMEM 3048 3060 POTENTIAL
 FT DOMAIN 266 1196 SER/THR-RICH.
 FT DOMAIN 1105 1241 GLY/SER-RICH.
 FT DOMAIN 2071 2120 GPS.
 FT DOMAIN 2182 2305 PLAT.
 SQ SEQUENCE 3178 AA: 344726 MW: 182394 (CD04666CD) CAC64:
 Query Match 3.98, Score 143, E-1; Length 3178;
 Best Local Similarity 20.68; Prod. No. 4.9;
 Matches 121, Conservative 61, Mismatches 288, Gaps 24:
 QY 53 APTVTAAGAGTFFVETPARTLQAKWQQA-KTAVASKEKTTTAGAAALAVAPAFITTV 122
 DB 1213 -TATVSSSTGPTTASNPSTITGSIENAIFFPAVAVSEPINHTLITIGTAALNYVFLVQ 1271
 QY 242 AFQFSMTSSSTLEPMFALEMPQYTESLSSTLPPHPS-----ATVTVSSTTV 296
 DB 1272 NPLSSAQQ--IFNPLPLVLK-----YFNSNTSSE--LDNSICE--LST 1307
 QY 123 FQVSEKCALEKQVNSSTPLSSVALGTHLVYAMLSPLGFLQSTNTHMAYASNY 182
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 DB 207 P-----TSTVTVTAMSTST-----TPSTSTTFSTSTSTSTSTSTST 434
 QY 183 AQVAVARATIRYP-----EVNAVGY-----ALISLW----- 212
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 DB 355 SLLQSSSSSTLSSSTSSLSLSLSTLSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 398
 QY 213 - POTTITTSVMNS-----TSTVNRKLVQGLASLAVIES----- 248
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 DB 399 LSTSTTTPST 458
 QY 243 EEEHYRQWESVETSEVALEKATKANNLECHISNVEYRHTTLEALNLLALELE 407
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 DB 459 TSTVSSSSST 504
 QY 608 EPLLENT-----ATVSPSSSTAPHPPLPACACIALPPLAALAVAPAFITTV 555
 DB 1213 -TATVSSSTGPTTASNPSTITGSIENAIFFPAVAVSEPINHTLITIGTAALNYVFLVQ 1271
 DB 505 SSSSIPSSSIASSSVSSSIASSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 561
 QY 456 -----TSTN-----GVEIGRGIAL-----TLNLAHL-----LGL 485
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 DB 562 EFTVSSVRLSS 621
 QY 569 LLSSAAALAVYFVVAAGARVAVYVVEALPFLAALHLLTAPVAVLSTN 446
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 QY 472 EFTVAVTAVLFAANLHLLPFWYVYAH-----VYVYSS 467
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 QY 446 QH-PQDPPT-----SPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 496
 DB 1308 FILSYRSGSSTTTLSPKSISSLSVW 1332
 QY 496 EFTVAVTAVTAVAVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 549
 DB 718 SEESTTSSSTTSTSTSTSSSSSSSTTADSTSTLSVQLD-----FLL 760

RESULT 12

YAG3_YEAST
 ID YAG3_YEAST STANDARD: PRT: 1422 AA.
 AC P39712;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1996 (Rel. 43, Last sequence update)
 DT 01-FEB-1996 (Rel. 43, Last annotation update)
 DE Hypothetical 149.1 kDa protein in P409-GDH's intergenic precursor.
 GN YAL063C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Eukaryota; Saccharomycetes; Saccharomycetaceae;
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN S288c / AP072;
 EX MEDLINE 9521563, PubMed 7719083
 RA Bockley H., Faller K., Bockley H., V. Bockley, Clark M.W., Fother N.


```

FT CONFLICT 1234 1234 N -> Y (IN REF. 2).
FT CONFLICT 1254 1254 V -> A (IN REF. 2).
FT CONFLICT 1282 1282 C -> R (IN REF. 2).
SQ SEQUENCE 1321 AA: 2EF47F823DB9H08 CRC64;

Query Match 3.8%; Score 129.5; DB 1: Length 1321;
Best local Similarity 22.0%; Pred.No. 2;
Matches 140, Conservative 63, Mismatches 231, Indels 201, Gaps 32;

QY 22 PQLFSSCHPRPPPSGSGSGSF-----WGRVDSQP-----FAIPYTHPTNPAPVPT 67
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 472 PTDMPRRPPGPKINGRYQQQPEPEGLQGMEASAOPTSEAAVNOMEP--PLAMAVT 529
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 AAAGAGPVROPARPLGSARFD---OACRPVAVSRRRPTTAGAA-----PLTAVAPA 116
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 530 EMIASG-QSRSP-----WADLTNEVDMPCAGS-----AGCKSSPEPLWPPNTVDPDS 575
QY 117 ---HTTPVPVVD-SRGAILR-KYNI-STPLTSSVATGINLVYAAPLS-----PLPL 166
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 576 ISCHSRAPVLELEKAEGPSARPATPDLFWSPLEATVSAPSPAPWEAFPVATSHDLMMAM 635
QY 167 QUNTINTHIM-----ATEASNYAGYRVARATIPYPPVPVNAVGGYALISFWPQTTTPT 220
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 636 LRQKEMLPHTPTISTEANKRVHGEATATAPPSPAATKV--YSLPLSL-----TPT 687
QY 221 SVMNSTITINVP!!VAPCTASPLVTPSEPIHYPNQOWPSVETSTVAEEEAYS--CLVM: 278
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 688 G-----QGEAMPTTPESPRADFPETGETSPAQVNKAHSSSSSPWPSVNP 732
QY 279 CIHGSUNSYNTPTGALGLIDFALEFRNLIDONTNTRVSRYSSTARRHLRGCADGT 338
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 733 NVAVGEVPTETATEPTGLRGI-----PGSES-----GVFDT 763
QY 339 AFLTTTAATREMKDLY-----FTSTNGVGEIGR-GIALT----- 371
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 764 AFSPTSGLOATVEVQNPWPSVYSKGLDASSPAGLGSFVLPKAVTPNLFVWVAIDEG 823
QY 372 -IPNIADTILGGIPTET--TSSAGGQLFYSPPVVSANGEPYKLYTSV-----ENAGQURK 424
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 824 PTVPMDSTVTPAPSDASGIWEPGSQVF--EEAESTITSPQVALDTSIVTPTLTLEQGEK 881
QY 425 IAIPIHDIDIGESRVVIQDYDQHPDRPTSP-----APSRPFVLRANUVI 471
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 882 VGVFAMSTLGSS-----SSQHPFPETQVETQCTSCASVPPHQSPLGKPAVP 929
QY 472 WLSLTAAEYDQSTYCSSTGPVYVSD-SVTLVNVATGAQAVARSLDMTKVTLDCRPLSTIO 530
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 930 PCTPTAASVGFSAVSSSSFTVPWDPSTLLPVTLGIE-----DPELEVLGASP----- 978
QY 531 QYKTFEVLPLRGKLSWF---AGTTKA--GYPN 560
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 979 -----GVESFEEVASGEFALPGPMN 1001
```

Search completed: August 5, 2003, 09:27:38
Job time : 31 secs

Genware version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

us-09-851-410a-8.std.rspt

```

121 PVPDVSRCAILRQYNLSTSPLTSSVATGTLNVLVAAPLSPLLPLODCTNTHIMATEAS 180
181 NYAUYRVARATIRYRPLVPNAVGGYALISFPWQDTTITPTSVIMNSITSTDVPLVQPGI 240
181 NYAUYRVARATIRYRPLVPNAVGGYALISFPWQDTTITPTSVIMNSITSTDVPLVQPGI 240
241 ASELVPSERLHYRNOGWRSVETSGVAEEATSGVLMCHGSLVNSYNTNPTVTGALGLL 300
241 ASELVPSERLHYRNOGWRSVETSGVAEEATSGVLMCHGSLVNSYNTNPTVTGALGLL 300
301 DFALFLEFRNLTPGNTNTRVSPYSSSTAPHRIPRGADGTAEITTTAATREMKKOLYFTSNG 360
301 DFALFLEFRNLTPGNTNTRVSPYSSSTAPHRIPRGADGTAEITTTAATREMKKOLYFTSNG 360
361 VGEIGRGIALILFNLADTLGSLTPELISAGGOLFYSRPVVSANGEPVTKLYTSVENAQ 420
361 VGEIGRGIALILFNLADTLGSLTPELISAGGOLFYSRPVVSANGEPVTKLYTSVENAQ 420
421 QDKGIAIPHDIDGSGSKVVLIGYGNQHEUQKPTSPAPSPFPFVSLKANDVWLSTAAEY 480
421 QDKGIAIPHDIDGSGSKVVLIGYGNQHEUQKPTSPAPSPFPFVSLKANDVWLSTAAEY 480
481 DQSTYSSSGVYVSDSVTLNVNATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
481 DQSTYSSSGVYVSDSVTLNVNATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
541 LRKGLSPWEAGTTRAGYVYNTTASQQLVENAAGHRVAISYITTSIGAGPVSISAVAV 600
541 LRKGLSPWEAGTTRAGYVYNTTASQQLVENAAGHRVAISYITTSIGAGPVSISAVAV 600
601 LAPHSALEETLIDYFARAITEGGRPPFPPGLQGPAPGSIVAHQLPKMKVKIKREL 660
601 LAPHSALEETLIDYFARAITEGGRPPFPPGLQGPAPGSIVAHQLPKMKVKIKREL 660

```

RESULT 2

```

091855 PRELIMINARY: PRI: 660 AA.
AC 091855;
DI 01-NOV-1998 (TREMBlrel. 08, Created)
DI 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DI 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93-Egypt;
RX MEDLINE=99105430; PubMed=4890424;
RA Tsarev S.A., Hino T.N., Gamalios P.J., Athar R.E., Meillet M.K.,
RA van Ouyck-Gandre H., Tonger C.F., Indis R.I.
KT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
R1 J. Med. Virol. 57:68-74(1999)
DR EMBL: AF051351; AAC35761.1;
DR InterPro: IPR004261; SP2;
DR Pfam: PF04014; SP2; 1;
SQ SEQUENCE 660 AA: 71040 MW: 044PF5EA7C492791 CRC94;

```

Query Match 99.18; Score 3385; ID 12; Length 660;
 Best Local Similarity 99.18; Pred. No. 5.4e-215;
 Matches 654; Conservatio 3; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MRPRPILLLLMLEMLPAPPGQPSGPPHPSGSGSPSPSPVPSAPFALPYLHPIN 50
DB 1 MRPRPILLLLMLEMLPAPPGQPSGPPHPSGSGSPSPSPVPSAPFALPYLHPIN 50
QY 61 PPAVNTAAACAPRVPQPARPLCSAWPVAQGVAVASRPPTTAAATLPAVAAMHPT 120
DB 61 PPAVNTAAACAPRVPQPARPLCSAWPVAQGVAVASRPPTTAAATLPAVAAMHPT 120
QY 121 PVTIVDSGCATIRKQVHSTSTETASVATGTLNVLVAAPLSPLLPLODCTNTHIMATEAS 180

```

```

121 PVPDVSRCAILRQYNLSTSPLTSSVATGTLNVLVAAPLSPLLPLODCTNTHIMATEAS 180
181 NYAUYRVARATIRYRPLVPNAVGGYALISFPWQDTTITPTSVIMNSITSTDVPLVQPGI 240
181 NYAUYRVARATIRYRPLVPNAVGGYALISFPWQDTTITPTSVIMNSITSTDVPLVQPGI 240
241 ASELVPSERLHYRNOGWRSVETSGVAEEATSGVLMCHGSLVNSYNTNPTVTGALGLL 300
241 ASELVPSERLHYRNOGWRSVETSGVAEEATSGVLMCHGSLVNSYNTNPTVTGALGLL 300
301 DFALFLEFRNLTPGNTNTRVSPYSSSTAPHRIPRGADGTAEITTTAATREMKKOLYFTSNG 360
301 DFALFLEFRNLTPGNTNTRVSPYSSSTAPHRIPRGADGTAEITTTAATREMKKOLYFTSNG 360
361 VGEIGRGIALILFNLADTLGSLTPELISAGGOLFYSRPVVSANGEPVTKLYTSVENAQ 420
361 VGEIGRGIALILFNLADTLGSLTPELISAGGOLFYSRPVVSANGEPVTKLYTSVENAQ 420
421 QDKGIAIPHDIDGSGSKVVLIGYGNQHEUQKPTSPAPSPFPFVSLKANDVWLSTAAEY 480
421 QDKGIAIPHDIDGSGSKVVLIGYGNQHEUQKPTSPAPSPFPFVSLKANDVWLSTAAEY 480
481 DQSTYSSSGVYVSDSVTLNVNATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
481 DQSTYSSSGVYVSDSVTLNVNATGAGAVARSLDWTKVTLGKRPSTTQQSKTFEVL 540
541 LRKGLSPWEAGTTRAGYVYNTTASQQLVENAAGHRVAISYITTSIGAGPVSISAVAV 600
541 LRKGLSPWEAGTTRAGYVYNTTASQQLVENAAGHRVAISYITTSIGAGPVSISAVAV 600
601 LAPHSALEETLIDYFARAITEGGRPPFPPGLQGPAPGSIVAHQLPKMKVKIKREL 660
601 LAPHSALEETLIDYFARAITEGGRPPFPPGLQGPAPGSIVAHQLPKMKVKIKREL 660

```

RESULT 3

```

09WQAO PRELIMINARY: PRI: 660 AA.
AC 09WQAO;
DI 01-NOV-1999 (TREMBlrel. 12, Created)
DI 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DI 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKL-90;
RX MEDLINE=99350000; PubMed=1042137;
RA Arankalle V.A., Taranjape S., Davison R.G., Barrett R.B.,
RA Walimbe A.M.;
KT "Phylogenetic analysis of hepatitis E virus isolates from India (1976-1999).";
R1 J. Gen. Virol. 80:1691-1700(1999);
DR EMBL: AF124407; AAD5493.1;
DR InterPro: IPR004261; SP2;
DR Pfam: PF04014; SP2; 1;
SQ SEQUENCE 660 AA: 71941 MW: 038720CF0A87B521 CRC94;

```

Query Match 99.97; Score 3383; ID 12; Length 660;
 Best Local Similarity 99.28; Pred. No. 7.4e-216;
 Matches 655; Conservatio 1; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MRPRPILLLLMLEMLPAPPGQPSGPPHPSGSGSPSPSPVPSAPFALPYLHPIN 50
DB 1 MRPRPILLLLMLEMLPAPPGQPSGPPHPSGSGSPSPSPVPSAPFALPYLHPIN 50
QY 61 PPAVNTAAACAPRVPQPARPLCSAWPVAQGVAVASRPPTTAAATLPAVAAMHPT 120
DB 61 PPAVNTAAACAPRVPQPARPLCSAWPVAQGVAVASRPPTTAAATLPAVAAMHPT 120
QY 121 PVTIVDSGCATIRKQVHSTSTETASVATGTLNVLVAAPLSPLLPLODCTNTHIMATEAS 180

```



```

10 4.1 JUKLHAIHIIHIGESPVVLTGYNQJHJLHAIHIGESPVVITKANAVIMSLTAAY 480
11
12 401 DESVSSSPGVVSVSVLVNVAFCAGAVAPSLWVKTLLCPPLSTIQVSKPFFVLP 540
13 0000000000000000000000000000000000000000000000000000000
14 401 POSTVSSSPGVVSVSVLVNVAFCAGAVAPSLWVKTLLCPPLSTIQVSKPFFVLP 540
15 000000000000000000000000000000000000000000000000000000
16 401 DEKLSFWAGTTPKATITVTRTALATLVNNAKHEVALITVLSLCAIVZAVAV 600
17 000000000000000000000000000000000000000000000000000000
18 401 LQKLSFWAGTTPKATITVTRTALATLVNNAKHEVALITVLSLCAIVZAVAV 600
19 000000000000000000000000000000000000000000000000000000
20 401 LAPHSALALLEDITLYPARAHITLHGFCEKPGIQLQVAPQSIVAEQLQKMKVKIKREL 660
21 000000000000000000000000000000000000000000000000000000
22 401 LAPHSALALLEDITLYPARAHITLHGFCEKPGIQLQVAPQSIVAEQLQKMKVKIKREL 660
23 000000000000000000000000000000000000000000000000000000

```

RESULT 7
 Q00419
 ID Q00419 PRELIMINARY: PRI: 660 AA
 AC Q00419
 DT 01 NOV 1996 (1-EMBLrel, 01, Created)
 DT 01 NOV 1996 (1-EMBLrel, 01, Last sequence update)
 DT 01 DEC 2001 (1-EMBLrel, 19, Last annotation update)
 DE Q0041, Q002 & Q003
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID 12461;
 RN 111
 RP SEQUENCE FROM N.A.
 RA Ron Ramo A., Sechurn J., Hyattarajan S.P., Mahanavalli B., Menon T.,
 RA "Q0041" amplification, cloning and sequence determination of a hepatitis
 RA E virus isolate from Madras, India."
 RI Submitted (001, 1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: X00441; Genbank: U00441; SFL:
 DR InterPro: IPR004261; SFL:
 DR Pfam: PF00014; SFL: 1;
 SQ SEQUENCE: 660 AA; 7060 MW; 1404099916098 CRC64;

Query Match 99.09; Score 1474; DB 12; Length 660;
 Best Local Similarity 99.09; Prod. No. 2,960,214;
 Matches 654; Conservation 6; Mismatches 6; Indels 6; Gaps 0;

```

QY 1 MKRPRLILLMLFPLMPLFAPHPGSPGRRKGRSGSGSGGFWGDRVDSDFEALPYTHPTN 60
DB 1 MKRPRLILLMLFPLMPLFAPHPGSPGRRKGRSGSGSGGFWGDRVDSDFEALPYTHPTN 60
QY 61 PFAPVIAAAGACHPVROPAPPLGASAWRQAGAGAVASPPRPPTAGAAPLTAVAPAHPTP 120
DB 61 PFAPVIAAAGACHPVROPAPPLGASAWRQAGAGAVASPPRPPTAGAAPLTAVAPAHPTP 120
QY 121 FVTVVDSGATLRRQYNLSPTSSVAVCTNIVYAAPLSLPLQKQNTNIMATPAS 180
DB 121 FVTVVDSGATLRRQYNLSPTSSVAVCTNIVYAAPLSLPLQKQNTNIMATPAS 180
QY 181 NYAGVYKAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
DB 181 NYAGVYKAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
QY 241 NYA-YGVAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
DB 241 NYA-YGVAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
QY 241 ALVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 400
DB 241 ALVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 400
QY 401 ASGVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 460
DB 401 ASGVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 460
QY 461 DEALLERPNLTNTNTRSPYSSIAHBLRKGAGTAEALITAAIRMKDLYFISINQ 480
DB 461 DEALLERPNLTNTNTRSPYSSIAHBLRKGAGTAEALITAAIRMKDLYFISINQ 480
QY 481 WETGKGIALTLFLNLAITLGLLTELSSAGQGLFYSRVVVSANGEPVVKLYTSVENAQ 420
DB 481 WETGKGIALTLFLNLAITLGLLTELSSAGQGLFYSRVVVSANGEPVVKLYTSVENAQ 420
QY 421 WETGKGIALTLFLNLAITLGLLTELSSAGQGLFYSRVVVSANGEPVVKLYTSVENAQ 420
DB 421 WETGKGIALTLFLNLAITLGLLTELSSAGQGLFYSRVVVSANGEPVVKLYTSVENAQ 420
QY 421 GKRGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480
DB 421 GKRGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480
QY 421 LFEGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480
DB 421 LFEGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480

```

Query Match 98.99; Score 1372; DB 12; Length 660;
 Best Local Similarity 98.99; Prod. No. 3,760,214;
 Matches 651; Conservation 4; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MKRPRLILLMLFPLMPLFAPHPGSPGRRKGRSGSGSGGFWGDRVDSDFEALPYTHPTN 60
DB 1 MKRPRLILLMLFPLMPLFAPHPGSPGRRKGRSGSGSGGFWGDRVDSDFEALPYTHPTN 60
QY 61 PFAPVIAAAGACHPVROPAPPLGASAWRQAGAGAVASPPRPPTAGAAPLTAVAPAHPTP 120
DB 61 PFAPVIAAAGACHPVROPAPPLGASAWRQAGAGAVASPPRPPTAGAAPLTAVAPAHPTP 120
QY 121 FVTVVDSGATLRRQYNLSPTSSVAVCTNIVYAAPLSLPLQKQNTNIMATPAS 180
DB 121 FVTVVDSGATLRRQYNLSPTSSVAVCTNIVYAAPLSLPLQKQNTNIMATPAS 180
QY 181 NYAGVYKAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
DB 181 NYAGVYKAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
QY 241 NYA-YGVAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
DB 241 NYA-YGVAPATITRKPIVPAVAGYALISFHWQTTTPTSVDMNSTITSDVHIVOPGI 240
QY 241 ALVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 400
DB 241 ALVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 400
QY 401 ASGVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 460
DB 401 ASGVAVPCHRGYENSGWVEVGVAVAGATSGVMGCHHSLVHSYNTNPTGASLLG 460
QY 461 DEALLERPNLTNTNTRSPYSSIAHBLRKGAGTAEALITAAIRMKDLYFISINQ 480
DB 461 DEALLERPNLTNTNTRSPYSSIAHBLRKGAGTAEALITAAIRMKDLYFISINQ 480
QY 481 WETGKGIALTLFLNLAITLGLLTELSSAGQGLFYSRVVVSANGEPVVKLYTSVENAQ 420
DB 481 WETGKGIALTLFLNLAITLGLLTELSSAGQGLFYSRVVVSANGEPVVKLYTSVENAQ 420
QY 421 GKRGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480
DB 421 GKRGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480
QY 421 LFEGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480
DB 421 LFEGLAIPGIDLSHGVVQGLVGNHDEHDEGATFAPSAPGVLSKANIWLMSLTAAAY 480

```

RESULT 8
 Q01878
 ID Q01878 PRELIMINARY: PRI: 660 AA
 AC Q01878
 DT 01 NOV 1996 (1-EMBLrel, 01, Created)
 DT 01 NOV 1996 (1-EMBLrel, 01, Last sequence update)
 DT 01 DEC 2001 (1-EMBLrel, 19, Last annotation update)
 DE Complete genome sequence.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID 12461;
 RN 111
 RP SEQUENCE FROM N.A.
 RA STRAIN UH1170;
 RA Isolated (WAY-5000) from the EMBL/Genbank/DBJ databases.
 RI EMBL: U11093; Genbank: U00467; SFL:
 DR InterPro: IPR004261; SFL:
 DR Pfam: PF00014; SFL: 1;
 SQ SEQUENCE: 660 AA; 70947 MW; 58123101426449 CRC64;


```

41 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
42 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
43 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560
44 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580
45 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
46 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620
47 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640
48 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660
49 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680
50 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700
51 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720
52 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740
53 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760
54 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780
55 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800
56 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820
57 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
58 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
59 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880
60 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900
61 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
62 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940
63 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960
64 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
65 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000
66 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
67 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040
68 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
69 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080
70 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
71 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120
72 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140
73 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160
74 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180
75 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
76 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220
77 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240
78 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260
79 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280
80 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300
81 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320
82 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
83 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360
84 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380
85 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400
86 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420
87 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440
88 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460
89 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480
90 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500
91 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520
92 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540
93 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560
94 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580
95 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600
96 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620
97 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640
98 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660
99 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680
100 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700
101 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720
102 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740
103 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760
104 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780
105 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800
106 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820
107 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840
108 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860
109 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880
110 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900
111 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920
112 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940
113 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960
114 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980
115 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000
116 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
117 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040
118 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060
119 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080
120 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100
121 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120
122 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140
123 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160
124 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180
125 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200
126 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220
127 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240
128 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260
129 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280
130 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300
131 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320
132 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340
133 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360
134 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380
135 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400
136 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420
137 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440
138 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460
139 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480
140 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500
141 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520
142 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540
143 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560
144 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580
145 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600
146 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620
147 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640
148 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660
149 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680
150 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700
151 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720
152 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740
153 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760
154 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780
155 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800
156 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820
157 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836 2837 2838 2839 2840
158 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860
159 2861 2862 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878 2879 2880
160 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900
161 2901 2902 2903 2904 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920
162 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934 2935 2936 2937 2938 2939 2940
163 2941 2942 2943 2944 2945 2946 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 295
```

```

Db 180 NYAGYVVARATIPYPPPIVNAVSGYALSISEWQTTPTPTSDVMNSVTSIGVPLVQPGI 239
QY 241 ASELVPSERLHYNNQWKPVSFVSGVAEPATSHVMLA'PHSLSVNSVNTPTTGAUGLI 400
Db 240 ASELVPSERLHYNNQWKPVSFVSGVAEPATSHVMLA'PHSLSVNSVNTPTTGAUGLI 299
QY 301 DFALELEFNLTICNTNTRVSRYSSTAPRHLRBCADGTAEITTAATREMKDLYFTSTNG 360
Db 300 DFALELEFNLTICNTNTRVSRYSSTAPRHLRBCADGTAEITTAATREMKDLYFTSTNG 359
QY 361 VGEIGRGIATLFLNADTLGGIPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 360 VGEIGRGIATLFLNADTLGGIPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 419
QY 421 QKKGTAIPNDIGESKWTQDYNQHEQLPPYSPAPSPPPSVLPAANDVWLSTAAEY 480
Db 420 QKKGTAIPNDIGESKWTQDYNQHEQLPPYSPAPSPPPSVLPAANDVWLSTAAEY 479
QY 481 DQSYGSGTGPVVSUSVTLVNAVATACAVARSLDWTIKVTLGGPSTTQQYSKTFPVLP 540
Db 480 DQSYGSGTGPVVSUSVTLVNAVATACAVARSLDWTIKVTLGGPSTTQQYSKTFPVLP 539
QY 541 LRKGLSPWEACTTKACYPYNNNTASDQILVENAAGHPVAISTYTTSLGAGPVVISAVAV 600
Db 540 LRKGLSPWEACTTKACYPYNNNTASDQILVENAAGHPVAISTYTTSLGAGPVVISAVAV 599
QY 601 LAPHSALALLEDTLDPARAKHIFDQF'PRKPKVGLDQ'APQSTVAFIQRIKMKVGTRELP 660
Db 600 LAPHSALALLEDTLDPARAKHIFDQF'PRKPKVGLDQ'APQSTVAFIQRIKMKVGTRELP 659

RESULT 13
Q68985
ID Q68985 PRELIMINARY: PRI: 660 AA.
AC Q68985:
DI 01-NOV-1996 (Tremblrel 01, created)
DI 01-NOV-1996 (Tremblrel 01, last sequence update)
DI 01-DEC-2001 (Tremblrel 10, last annotation update)
DE GRF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN HYDEPARAD.
RX MEDLINE=96087204 PubMed=3-8567903.
RA Panda S.K., Nanda S.K., Zafarullah M., Ansari I.H., Ordener M.H.,
Jameel S.;
RI "An Indian strain of hepatitis E virus (HEV). Cloning, sequencing, and
expression of structural region and antibody responses in sera from
individuals from an area of high-level HPE endemicity";
RJ J. Clin. Microbiol. 33:2653-2659(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Ansari I.H., Nanda S.K., Dargopal H., Jameel S., Panda S.K.;
RI "Translational analysis of hepatitis E virus (HEV) genome";
RJ Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR FMBL: 022512; AAA97366.1; -
DR EMBL: AF042930; AAC29936.1; -
DR InterPro: IPR004261; SP2.
DR Pfam: PF03014; SP2; 1
SQ SEQUENCE: 660 AA: 70936 MW: 948566 P04362 P04362 C0004.

Query Match 97.94; Score 4336; E: 1e-12; Length 660;
Best Local Similarity 98.08; Pred. No. 9.3e-012;
Matches 647; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 MGPRIILLLMETPMI:PATPQCPGSPRRP'RRSGSGGCPWQWQVNSQPPALPYHTPTN 60
Db 1 MGPRIILLLMETPMI:PATPQCPGSPRRP'RRSGSGGCPWQWQVNSQPPALPYHTPTN 60
QY 61 PFAPDVTAAAGACGAPPPVQPPAPPLASAMPQCAQAPAVASPPPTPTTA'CAAPLTAVATAHET 120

```

```

Db 61 PFAPDVTAAAGACGAPPPVQPPAPPLASAMPQCAQAPAVASPPPTPTTA'CAAPLTAVATAHET 120
QY 121 PVPIVUSKGAILLKROVNLSTPLTSSVATGNIIVIAAPLSPLLPLAQGNIHIMAEAS 180
Db 121 PVPIVUSKGAILLKROVNLSTPLTSSVATGNIIVIAAPLSPLLPLAQGNIHIMAEAS 180
QY 181 NYAGYVVARATIPYPPPIVNAVSGYALSISEWQTTPTPTSDVMNSVTSIGVPLVQPGI 240
Db 181 NYAGYVVARATIPYPPPIVNAVSGYALSISEWQTTPTPTSDVMNSVTSIGVPLVQPGI 240
QY 241 ASELVPSERLHYNNQWKPVSFVSGVAEPATSHVMLA'PHSLSVNSVNTPTTGAUGLI 400
Db 241 ASELVPSERLHYNNQWKPVSFVSGVAEPATSHVMLA'PHSLSVNSVNTPTTGAUGLI 400
QY 301 DFALELEFNLTICNTNTRVSRYSSTAPRHLRBCADGTAEITTAATREMKDLYFTSTNG 360
Db 301 DFALELEFNLTICNTNTRVSRYSSTAPRHLRBCADGTAEITTAATREMKDLYFTSTNG 360
QY 361 VGEIGRGIATLFLNADTLGGIPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIATLFLNADTLGGIPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QKKGTAIPNDIGESKWTQDYNQHEQLPPYSPAPSPPPSVLPAANDVWLSTAAEY 480
Db 421 QKKGTAIPNDIGESKWTQDYNQHEQLPPYSPAPSPPPSVLPAANDVWLSTAAEY 480
QY 481 DQSYGSGTGPVVSUSVTLVNAVATACAVARSLDWTIKVTLGGPSTTQQYSKTFPVLP 540
Db 481 DQSYGSGTGPVVSUSVTLVNAVATACAVARSLDWTIKVTLGGPSTTQQYSKTFPVLP 540
QY 541 LRKGLSPWEACTTKACYPYNNNTASDQILVENAAGHPVAISTYTTSLGAGPVVISAVAV 600
Db 541 LRKGLSPWEACTTKACYPYNNNTASDQILVENAAGHPVAISTYTTSLGAGPVVISAVAV 600
QY 601 LAPHSALALLEDTLDPARAKHIFDQF'PRKPKVGLDQ'APQSTVAFIQRIKMKVGTRELP 660
Db 601 LAPHSALALLEDTLDPARAKHIFDQF'PRKPKVGLDQ'APQSTVAFIQRIKMKVGTRELP 660

RESULT 14
Q68986
ID Q68986 PRELIMINARY: PRI: 674 AA.
AC Q68986:
DI 01-OCT-2002 (Tremblrel 22, created)
DI 01-OCT-2002 (Tremblrel 22, last sequence update)
DI 01-MAR-2003 (Tremblrel 23, last annotation update)
DE OYE2 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN JAK-Sait.
RA Mishiro S.;
RI "Serotype 4 (H.V. 2001) of the HBeAg-negative HPE isolates";
RJ J. Infect. Dis. 185:1442-1445(2002).
DR EMBL: AF044915; BAB9667.1; -
DR InterPro: IPR004261; SP2.
DR Pfam: PF03014; SP2; 1
SQ SEQUENCE: 674 AA: 73563 MW: 9016612 P04265 C0004.

Query Match 93.39; Score 4070; E: 1e-12; Length 674;
Best Local Similarity 92.18; Pred. No. 2.4e-011;
Matches 609; Conservative 100; Mismatches 27; Indels 0; Gaps 0;

```


[illegible]

com unipol (c)

Search completed: Yes
Job time : 13 sec

